

The Multivariate Kyle model: More is different

L. C. García del Molino¹, I. Mastromatteo¹, M. Benzaquen^{2,1}, and J.-P. Bouchaud¹

¹*Capital Fund Management, 23 rue de l'Université, 75007, Paris, France*

²*Ladhyx and Department of Economics, UMR CNRS 7646, Ecole polytechnique, 91128 Palaiseau Cedex, France*

December 21, 2018

Abstract

We reconsider the multivariate Kyle model in a risk-neutral setting with a single, perfectly informed rational insider and a rational competitive market maker, setting the price of n correlated securities. We prove the unicity of a symmetric, positive definite solution for the impact matrix and provide insights on its interpretation. We explore its implications from the perspective of empirical market microstructure, and argue that it provides a sensible inference procedure to cure some pathologies encountered in recent attempts to calibrate cross-impact matrices. As an illustration, we determine the empirical cross impact matrix of US Treasuries, and compare the results with recent alternative calibration methods.

Keywords: Microstructure, Impact, Multivariate, Price formation

Contents

1	Introduction	2
2	The Multivariate Kyle Model	3
2.1	The Model	3
2.2	Observables	3
3	Equilibrium strategies	4
3.1	Linear equilibrium	5
3.2	Price and volume covariances of the equilibrium solution	7
3.3	Utilities and competitive market making	8
4	Interpretation	10
4.1	Whitening of prices and volumes	10
4.2	Basis of prices and basis of volumes	11
5	Implications	13
5.1	From theory to data: empirical averages and loss function	13
5.2	Cross-impact estimators	14
5.2.1	Maximum Likelihood Estimation	14
5.2.2	EigenLiquidity model	15
5.2.3	Kyle estimator	16
5.3	A comparison of recipes	18
5.3.1	Application to real data	18
5.3.2	Synthetic data	19
6	Conclusion	22

A	Solution of the Multivariate Kyle model	25
A.1	Optimality of the Informed Trader	25
A.2	Optimality of the Market Maker	25
A.3	Symmetry of the linear equilibrium	26
A.4	Explicit form of the linear equilibrium	27
B	Estimation	27
B.1	The loss χ^2	27
B.2	Maximum Likelihood Estimator	28
B.3	EigenLiquidity Estimator	28
B.4	Kyle Estimator	29
B.5	Fabricating synthetic covariance matrices inspired from real data	29

1 Introduction

Understanding market impact – the mechanism through which trades tend to push prices – is with no doubt a venture of paramount importance. From the theoretical point of view, market impact lies at the very heart of price formation in financial markets. From the practitioners perspective, market impact is often at the origin of non negligible trading costs that need to be controlled to optimise execution strategies. In the past decades most of the literature has focused on the price impact of single products (for a recent review, see [BBDG18]), with no regard of inter-asset interactions. However, many market participants trade large portfolios that combine hundreds or thousands of assets. Thus addressing the matter of inter-asset price impact, coined *cross-impact*, is one of great interest both fundamentally and practically. Recent empirical studies show evidence of significant cross-impact effects in stock markets [HS01, PV13, WSG16, BMEB17]. A recurrent issue is that of empirical noise when it comes to large matrix estimation. It is thus important to continue to search for good statistical priors in order to help “clean” these large dimensional estimators.

What should one expect from theoretical economics on this matter? Which empirical observations should be considered unusual? Here we focus on how to make sense of the empirical observations accumulated over the past few years in the most orthodox setting: the classic Kyle model [Kyl85], extended to a multi-asset framework. Our aim is to gain insight into (i) how information is diffused into prices (cross-sectionally) within the Kyle setting and (ii) how one can use such results to regularise the very noisy regressions that arise in empirical cross-impact analysis. The multivariate Kyle model was first considered in [CK94], in a very general setting with n assets and m partially informed traders. The generality of the model is such that only a partial analysis of the solution is possible; furthermore, it involves an object that cannot be measured directly on empirical data, namely the correlation of the order flow of informed traders. Here we reconsider the problem with the issue of empirical calibration in mind. We focus on the particular case $m = 1$ and provide a proof that the linear equilibrium is unique and characterized by a symmetric, positive definite (SPD) impact matrix. This solution provides an explicit recipe to infer cross impact from (cleaned) order flow and return correlation matrices, that we compare to other natural recipes, such as Maximum Likelihood Estimators (MLE) [BMEB17] or the recently proposed “EigenLiquidity” Model (ELM) [MBEB17].

The paper is organised as follows. In sections 2 and 3 we introduce the multivariate Kyle model and provide the equilibrium strategies, mostly building upon the work in [CK94] but also providing new results. In section 4 we study in detail the mathematical properties of the solution. In section 5, we introduce an impact estimator based on the equilibrium strategy of the market maker and compare it to other usual estimators both from the theoretical and empirical point of view. Along the text we present several examples intended to provide intuition behind the main results. In particular it is often interesting to confront the results of the multivariate and univariate models.

In all the following bold uppercase symbols denote matrices, bold lower cases denote vectors and light lower cases denote scalars.

2 The Multivariate Kyle Model

In this section, we present in detail the multivariate Kyle setting and define the observables that would allow one to calibrate the model using empirical observables.

2.1 The Model

Consider a single-period economy where three representative agents trade n instruments. The agents are: an informed trader (IT) who has perfect information about the future prices \mathbf{v} , a noise trader (NT) that trades in absence of any information due to exogenous reasons, and a competitive market maker (MM) that has the role of enforcing price efficiency. The dynamics of the model is set by the following rules.

1. A fundamental price \mathbf{v} is sampled from a Gaussian distribution $\mathbf{v} \sim \mathcal{N}(\mathbf{p}_0, \mathbf{\Sigma}_0)$ where $\mathbf{\Sigma}_0$ is SPD. Only the IT knows the value of \mathbf{v} in advance, while \mathbf{p}_0 and $\mathbf{\Sigma}_0$ are common knowledge. We denote the price deviation from its mean as $\Delta \mathbf{v} := \mathbf{v} - \mathbf{p}_0$.
2. The IT and NT place simultaneously their orders of sizes \mathbf{x} and \mathbf{u} respectively. The bids of the NT \mathbf{u} are sampled from a Gaussian distribution $\mathbf{u} \sim \mathcal{N}(0, \mathbf{\Omega})$ independent of the fundamental price, where $\mathbf{\Omega}$ is an invertible matrix.
3. The MM clears the excess demand \mathbf{y} at a clearing price \mathbf{p} based on the total observed order imbalance $\mathbf{y} = \mathbf{x} + \mathbf{u}$, that allows him/her to form the best estimation of the fundamental prices. These fundamental prices are then revealed.

The quantity \mathbf{x} requested by the *risk-neutral* IT is such that he maximises the expectation of his utility function:

$$\mathcal{U}_{IT}(\mathbf{x}, \mathbf{p}) = \mathbf{x}^\top (\mathbf{v} - \mathbf{p}). \quad (1)$$

Note that \mathcal{U}_{IT} does not contain any risk penalty. While the introduction of such a penalty would affect some of the conclusions below, we decide to leave this interesting issue for future work. Consistent with the assumption that market making is competitive, the MM sets a price that matches in expectation \mathbf{v} given the available public information, namely the total order imbalance \mathbf{y} :

$$\mathbf{p} = \mathbb{E}[\mathbf{v} | \mathbf{y}], \quad (2)$$

where $\mathbb{E}[\cdot]$ denotes average with respect to the distribution of \mathbf{v} and \mathbf{u} .

Clearly, the above setting is highly stylized and, on many counts, unrealistic.¹ Still, this model is able to capture some of the essential ingredients of a reasonable price-formation process: the information owned by ITs gets encoded into a trading order imbalance \mathbf{x} polluted by a noise \mathbf{u} . In a competitive regime, MMs are expected to decode the information contained in the total order imbalance $\mathbf{y} = \mathbf{x} + \mathbf{u}$, in order to provide the best possible prediction of the fundamental price \mathbf{v} . This mechanically induces market impact: because the order imbalance \mathbf{y} is correlated with the fundamental price, the traded price \mathbf{p} will also display a correlation with the order imbalance \mathbf{y} , by that providing a sensible measure of market impact.

2.2 Observables

In order to gain insight on the implications of this model and in order to make testable predictions, one is required to provide some metrics that can be compared against market data. Luckily enough, the Gaussian nature of the setup allows one to only consider first- and second-order statistics (means and covariances) of

¹For example, the order flow in a Kyle setting has no temporal auto-correlations, whereas it is well known that the empirical order flow has long memory. See [BBDG18] and [BMEB17] for a recent discussion in the multivariate context.

prices and volumes in order to completely characterize the behavior of the model. Naturally, the fundamental parameters defining the model ($\boldsymbol{\Omega}, \boldsymbol{\Sigma}_0$) are not directly observable, and need to be inferred from the statistics of trades prices \boldsymbol{p} , and of volumes \boldsymbol{y} , which are the only physical observables of the model.

Prices Due to the price efficiency condition (see Eq. (2)), the average traded price is equal to the average of the fundamental price, itself equal (for consistency) to the initial price:

$$\mathbb{E}[\boldsymbol{p}] = \mathbb{E}[\boldsymbol{v}] = \boldsymbol{p}_0. \quad (3)$$

However, the covariance of the traded price and that of the fundamental price have no reason to coincide:

$$\boldsymbol{\Sigma} = \mathbb{C}[\boldsymbol{p}, \boldsymbol{p}] = \mathbb{E}[(\boldsymbol{p} - \boldsymbol{p}_0)(\boldsymbol{p} - \boldsymbol{p}_0)^\top], \quad (4)$$

$$\boldsymbol{\Sigma}_0 = \mathbb{C}[\boldsymbol{v}, \boldsymbol{v}] = \mathbb{E}[(\boldsymbol{v} - \boldsymbol{p}_0)(\boldsymbol{v} - \boldsymbol{p}_0)^\top]. \quad (5)$$

Further down we show that, at equilibrium, $\boldsymbol{\Sigma}$ and $\boldsymbol{\Sigma}_0$ are exactly proportional (see Eq. (24)).

Volumes Due to the non-informed nature of the NT, the average order imbalance is fixed by the bias introduced by the IT, so that:

$$\mathbb{E}[\boldsymbol{y}] = \mathbb{E}[\boldsymbol{x}] := \boldsymbol{x}_0. \quad (6)$$

The relation between the portion of volume covariance due to the NT and the one due to the IT is more subtle, and is given by:

$$\boldsymbol{\Omega}^d = \mathbb{C}[\boldsymbol{y}, \boldsymbol{y}] = \mathbb{C}[\boldsymbol{x}, \boldsymbol{x}] + \boldsymbol{\Omega}, \quad (7)$$

where we have introduced what we have coined the *dressed volume covariance* $\boldsymbol{\Omega}^d$, which is the physically observable quantity. Correspondingly, the *bare* volume covariance (i.e. not dressed by the noise contribution $\boldsymbol{\Omega}$) is $\mathbb{C}[\boldsymbol{x}, \boldsymbol{x}]$.

Response Ultimately, we are interested in characterising the expected price changes conditionally to a given trade imbalance: the response function, directly related to price impact. Within this model, the physical observables measuring such quantities, that we coin *dressed responses*, are:

$$\boldsymbol{R}^d = \mathbb{E}[(\boldsymbol{p} - \boldsymbol{p}_0)\boldsymbol{y}^\top] = \mathbb{C}[\boldsymbol{p}, \boldsymbol{y}], \quad (8)$$

$$\boldsymbol{R}_v^d = \mathbb{E}[(\boldsymbol{v} - \boldsymbol{p}_0)\boldsymbol{y}^\top] = \mathbb{C}[\boldsymbol{v}, \boldsymbol{y}]. \quad (9)$$

Due to the absence of correlations between the non-informed order imbalance \boldsymbol{u} and \boldsymbol{v} , one can relate the dressed responses with the responses with respect to \boldsymbol{x} : \boldsymbol{R} and \boldsymbol{R}_v , that we coin *bare responses*, through

$$\boldsymbol{R} = \mathbb{E}[(\boldsymbol{p} - \boldsymbol{p}_0)\boldsymbol{x}^\top] = \mathbb{C}[\boldsymbol{p}, \boldsymbol{x}] = \boldsymbol{R}^d - \mathbb{C}[\boldsymbol{p}, \boldsymbol{u}], \quad (10)$$

$$\boldsymbol{R}_v = \mathbb{E}[(\boldsymbol{v} - \boldsymbol{p}_0)\boldsymbol{x}^\top] = \mathbb{C}[\boldsymbol{v}, \boldsymbol{x}] = \boldsymbol{R}_v^d. \quad (11)$$

Interestingly, while the fundamental prices are insensitive to the level of noise trading, the traded prices incorporate some degree of mispricing due to the spurious correlations between the total order imbalance \boldsymbol{y} and the fundamental prices \boldsymbol{v} .

3 Equilibrium strategies

In this section we characterise the equilibrium strategies of the multivariate Kyle model. Although most of the results can be inferred from the seminal work of Caballé and Krishnan [CK94], we believe it is useful to provide a streamlined version of our own proofs, which we present in a more pedagogical and in some cases more compact form and within which several theoretical issues and empirical implications can be discussed explicitly.

3.1 Linear equilibrium

Our main result is the existence of a unique linear equilibrium as defined below, whose structure can be expressed analytically.

Definition 1 (Linear equilibrium). *By linear equilibrium, we mean a set of strategies in which the MM fixes the traded prices \mathbf{p} and the IT fixes their bid \mathbf{x} by means of linear rules*

$$\mathbf{p} = \boldsymbol{\mu} + \boldsymbol{\Lambda} \mathbf{y}, \quad (12)$$

$$\mathbf{x} = \boldsymbol{\alpha} + \mathbf{B} \mathbf{v}, \quad (13)$$

such that the two following conditions are satisfied:

1. Profit maximization: for all alternative strategies with $\mathbf{x}' \neq \mathbf{x}$,

$$\mathbb{E}[\mathcal{U}_{IT}(\mathbf{x}, \mathbf{p}) | \mathbf{v}] > \mathbb{E}[\mathcal{U}_{IT}(\mathbf{x}', \mathbf{p}) | \mathbf{v}].$$

2. Price efficiency: The price \mathbf{p} satisfies (2).

Based on the assumption of a linear strategy for the MM as given in Eq. (12), we obtain the following results (proofs are provided in App. A).

Proposition 1. *Imposing the linear pricing rule (12) for the MM implies that a rational IT will also set the order imbalance \mathbf{x} as a linear function of the imbalance, with:*

$$\mathbf{x} = \frac{1}{2} \boldsymbol{\Lambda}_S^{-1} (\mathbf{v} - \boldsymbol{\mu}), \quad (14)$$

where $\boldsymbol{\Lambda}_S$ denotes the symmetric part of $\boldsymbol{\Lambda}$. Furthermore, the profit maximization condition for the IT's strategy implies that $\boldsymbol{\Lambda}_S$ has to be positive-definite (PD).

Eq. (12) shows that $\boldsymbol{\Lambda}$ plays the role of adjusting the traded price level \mathbf{p} proportionally to the imbalance \mathbf{y} . Proposition 1 shows that $\boldsymbol{\Lambda}$ also plays the role of setting the order imbalance from the informed trader \mathbf{x} given the knowledge of the fundamental price \mathbf{v} . Moreover, Eq. (14) together with Eq. (12) imply that both the order imbalance \mathbf{y} and the traded price \mathbf{p} are normally distributed random variables, due to stability of the Gaussian distribution under convolution. This last property is at the core of the following proposition.

Proposition 2. *Assuming a linear strategy for the IT as in Eq. (14) implies that the parameters of the pricing rule of Eq. (12) for the MM are given by:*

$$\begin{aligned} \boldsymbol{\mu} &= \mathbf{p}_0 - \boldsymbol{\Lambda} \mathbf{y}_0, \\ \boldsymbol{\Lambda} &= \mathbf{R}_v^d (\boldsymbol{\Omega}^d)^{-1}. \end{aligned} \quad (15)$$

Note that \mathbf{y}_0 , \mathbf{R}_v^d and $\boldsymbol{\Omega}^d$ depend on \mathbf{x} , and therefore of $\boldsymbol{\Lambda}_S$. Substituting (15) into Eq. (14) allows to close the system and find an equation for $\boldsymbol{\Lambda}$. Note however that the as obtained system has many possible solutions. To constrain the latter we must impose the profit maximization condition introduced in Proposition 1.

Proposition 3. *Assume that there exists a solution to the utility maximization problem of the IT of the form given by Eq. (14), and to the pricing rule for the MM given by Eqs. (12) and (15). Then, the profit maximization condition that $\boldsymbol{\Lambda}_S$ has to be PD implies that $\boldsymbol{\Lambda}$ is symmetric and satisfies the equation:*

$$\frac{1}{4} \boldsymbol{\Sigma}_0 = \boldsymbol{\Lambda} \boldsymbol{\Omega} \boldsymbol{\Lambda}. \quad (16)$$

The symmetry of $\boldsymbol{\Lambda}$ leads to a unique solution that can be expressed in terms of the parameters of the problem (i.e. $\boldsymbol{\Sigma}_0$ and $\boldsymbol{\Omega}$). Denoting by $\sqrt{\mathbf{Y}}$ the unique (see Lemma 1 in App. A.3) PD solution of the matrix equation $\mathbf{X} \mathbf{X} = \mathbf{Y}$ for a SPD \mathbf{Y} we introduce the following theorem:

Theorem 1 (Existence and unicity of the linear equilibrium). *There exists a unique linear equilibrium given by the strategies (12) and (14) where:*

$$\mathbf{\Lambda} = \frac{1}{2}\mathcal{R}^{-1}\sqrt{\mathcal{R}\Sigma_0\mathcal{L}}\mathcal{L}^{-1}, \quad (17)$$

$$\mathbf{y}_0 = 0, \quad (18)$$

$$\boldsymbol{\mu} = \mathbf{p}_0. \quad (19)$$

Here \mathcal{L}, \mathcal{R} are a factorisation of the matrix $\boldsymbol{\Omega} = \mathcal{L}\mathcal{R}$ satisfying $\mathcal{L} = \mathcal{R}^\top$.

Note that the existence result is a special case of Proposition 3.1 from Ref. [CK94], specialised to the case of a single IT with perfect information. Nonetheless, the absence of other equilibria with non-symmetric $\mathbf{\Lambda}$ lacked a published proof (although a similar discussion can be found in an unpublished paper by Caballé and Krishnan [CK⁺90]). Because the symmetry property of $\mathbf{\Lambda}$ has crucial consequences on both price formation and pricing (see discussion in Sections 4 and 5) we believe that ruling out the existence of non-symmetric equilibria is an important step in the analysis of the multivariate Kyle model. This is the main novel theoretical contribution of the present paper.

Example 1 (Solution of the univariate Kyle model). *The specialisation of Theorem 1 to the $n = 1$ case in which only one asset is traded yields the well-known solution derived in [Kyl85] (light lower case symbols are scalar versions of the bold upper case and bold lower case ones):*

$$\lambda = \frac{1}{2}\sqrt{\frac{\sigma_0}{\omega}}, \quad (20)$$

indicating that the constant of proportionality between price and imbalances scales with the amount of price fluctuations $\sqrt{\sigma_0}$ and is inversely proportional to the typical fluctuations of the noise $\sqrt{\omega}$. Intuitively, the larger price deviations the rational MM expects, the more weight he should give to volume imbalances in order to forecast the fundamental price to the imbalances. On the other hand, the more noise there is in the system, the less the volume signal is reliable, so that the traded price is closer to the uninformed prior p_0 .

It is important to note that Eq. (16) alone does not imply symmetry of $\mathbf{\Lambda}$, as it is necessary to further impose PDness in order to obtain symmetry. In fact, there are symmetric solutions to (16) that are not PD. These solutions lead to efficient traded prices but they do not optimise the IT's utility.

Example 2 (Saddle point solutions). *Let \mathbf{x} be a strategy of the form given in (14) and let $\boldsymbol{\mu} = \mathbf{p}_0$ and let $\mathbf{\Lambda}$ be a symmetric but not PD solution of (16). By construction, the gradient of $\mathbb{E}[\mathcal{U}_{IT}]$ with respect to \mathbf{x} at this point is zero, satisfying the first order condition for the IT, but this does not guarantee that it is a maximum. We now take an arbitrary perturbation of the strategy $\mathbf{x} + \delta\mathbf{x}$ and compute the difference in utility between the two strategies:*

$$\begin{aligned} \delta\mathbb{E}[\mathcal{U}_{IT}] &= \mathbb{E}[\mathcal{U}_{IT}](\mathbf{x}) - \mathbb{E}[\mathcal{U}_{IT}](\mathbf{x} + \delta\mathbf{x}) \\ &= \mathbb{E}[2\mathbf{x}^\top \mathbf{\Lambda} \delta\mathbf{x} + \delta\mathbf{x}^\top \mathbf{\Lambda} \delta\mathbf{x} - \delta\mathbf{x}^\top (\mathbf{v} - \mathbf{p}_0)] \\ &= \delta\mathbf{x}^\top \mathbf{\Lambda} \delta\mathbf{x}. \end{aligned}$$

One can see that $\delta\mathbb{E}[\mathcal{U}_{IT}]$ is always positive for arbitrary $\delta\mathbf{x}$ if and only if $\mathbf{\Lambda}$ is PD. Without the second order condition the IT could find better strategies than \mathbf{x} meaning that the solution that corresponds to that $\mathbf{\Lambda}$ is not an equilibrium but a saddle point. In fact, when $\mathbf{\Lambda}$ is not PD, $\mathbb{E}[\mathcal{U}_{IT}]$ can be arbitrarily large.

In such a case, the IT would trade arbitrarily large quantities of some linear combinations of assets, and no equilibrium would be possible. Imposing a PD impact matrix $\mathbf{\Lambda}$ allows the MM to fix the optimal strategy of the IT, which otherwise would be undetermined – at least in the absence of a further risk penalty in the IT utility function.

3.2 Price and volume covariances of the equilibrium solution

We now investigate what are the testable implications of the linear equilibrium described above. Here, we provide the relations between the observables that one would measure in such a situation.

Proposition 4 (Camouflage). *In the equilibrium, the expectation of the dressed order imbalance is equal to:*

$$\mathbf{y}_0 = \mathbb{E}[\mathbf{y}] = \mathbb{E}[\mathbf{x}] = 0. \quad (21)$$

The IT order imbalance covariance can be expressed as:

$$\mathbb{C}[\mathbf{x}, \mathbf{x}] = \mathbf{\Omega}, \quad (22)$$

implying that:

$$\mathbf{\Omega}^d = 2\mathbf{\Omega}. \quad (23)$$

Proof. The results in (21) and (22) are obtained by directly evaluating the expectations:

$$\begin{aligned} \mathbb{E}[\mathbf{x}] &= \frac{1}{2}\mathbf{\Lambda}^{-1}\mathbb{E}[\mathbf{v} - \mathbf{p}_0] = 0, \\ \mathbb{C}[\mathbf{x}, \mathbf{x}] &= \frac{1}{4}\mathbf{\Lambda}^{-1}\mathbb{E}[(\mathbf{v} - \mathbf{p}_0)(\mathbf{v} - \mathbf{p}_0)^\top]\mathbf{\Lambda}^{-1} = \frac{1}{4}\mathbf{\Lambda}^{-1}\mathbf{\Sigma}_0\mathbf{\Lambda}^{-1} = \mathbf{\Omega}. \end{aligned}$$

The last result follows trivially from (22) and the definition of $\mathbf{\Omega}^d$. □

This result implies that, in addition to matching the trading level of the NT in order to conceal his information (as obtained in the one-dimensional case), the IT is required to match the *directions* in which the NT trades (i.e. the eigenvectors of $\mathbf{\Omega}$) such that the trading behaviour of the IT is indistinguishable from that of the NT. An even more interesting consequence of this behaviour is given in the next proposition.

Proposition 5 (Information diffusion). *The fundamental price covariance and the traded price covariance in the equilibrium are related by:*

$$\mathbf{\Sigma} = \frac{1}{2}\mathbf{\Sigma}_0. \quad (24)$$

Furthermore, the residual information is

$$\begin{aligned} \mathbb{C}[\mathbf{v}, \mathbf{v}|\mathbf{p}] &= \mathbb{C}[\mathbf{v}, \mathbf{v}] - \mathbb{C}[\mathbf{v}, \mathbf{p}]\mathbb{C}[\mathbf{p}, \mathbf{p}]^{-1}\mathbb{C}[\mathbf{p}, \mathbf{v}] \\ &= \mathbf{\Sigma}_0 - \mathbf{B}\mathbf{\Lambda}\mathbf{\Sigma}_0\mathbf{\Sigma}^{-1}\mathbf{\Sigma}_0\mathbf{\Lambda}^\top\mathbf{B}^\top \end{aligned} \quad (25)$$

which at equilibrium reads

$$\mathbb{C}[\mathbf{v}, \mathbf{v}|\mathbf{p}] = \frac{1}{2}\mathbf{\Sigma}_0. \quad (26)$$

Proof. The traded price covariance is obtained as:

$$\mathbf{\Sigma} = \mathbb{C}[\mathbf{p}, \mathbf{p}] = \mathbf{\Lambda}\mathbf{\Omega}^d\mathbf{\Lambda}^\top = 2\mathbf{\Lambda}\mathbf{\Omega}\mathbf{\Lambda} = \frac{1}{2}\mathbf{\Sigma}_0, \quad (27)$$

where the last equality comes from Eq. (73) in App. A.3. To find the residual information one needs to compute the covariance between traded prices and fundamental prices:

$$\mathbb{C}[\mathbf{v}, \mathbf{p}] = \mathbf{B}\mathbf{\Lambda}\mathbb{C}[\mathbf{v}, \mathbf{v}] = \mathbf{B}\mathbf{\Lambda}\mathbf{\Sigma}_0.$$

□

While the fact that the two covariances are related is expected, the fact that they are proportional to one another is non-trivial. Furthermore (27) implies that the role of the impact matrix $\mathbf{\Lambda}$ is to “rotate” the fluctuations of the volume in the direction of the fluctuations of the fundamental price. The intuition behind this behaviour is that due to camouflage the IT will trade in the same direction as the NT. Hence, a rational MM trying to enforce price efficiency will be required to convert fluctuations along the principal components of $\mathbf{\Omega}$ into forecasts of price fluctuations along the principal components of $\mathbf{\Sigma}_0$, thus matching the expected covariance $\mathbf{\Sigma}_0$. The fact that the proportionality constant is equal to 1/2 (only half of the information is revealed) is of minor importance, and is a consequence of the single-period feature of the model. Extending the Kyle framework to multiple time steps would turn the constant factor 1/2 into a time-varying function expressing the rate at which information is incorporated into prices (see the original Kyle paper [Kyl85] for the single asset case and [Vit12, Las04] for respectively the discrete and continuous multi-asset case).

Finally, it is interesting to characterise the responses in the context of a multivariate Kyle model.

Proposition 6 (Responses). *The propositions above and Eq. (10) imply that*

$$\mathbf{R}_v = \mathbf{R}_v^d = \mathbf{R}^d = 2\mathbf{R} = \frac{1}{2}\mathbf{\Sigma}_0\mathbf{\Lambda}^{-1}, \quad (28)$$

so that the responses are uniquely determined by $\mathbf{\Sigma}_0$ and $\mathbf{\Lambda}$.

Proof. The equilibrium values of the responses are found by plugging the equilibrium strategies on their definitions:

$$\begin{aligned} \mathbf{R}_v^d = \mathbf{R}_v &= \mathbb{E}[(\mathbf{v} - \mathbf{p}_0)\mathbf{x}^\top] = \frac{1}{2}\mathbb{E}[(\mathbf{v} - \mathbf{p}_0)(\mathbf{v} - \mathbf{p}_0)^\top]\mathbf{\Lambda}^{-1} = \frac{1}{2}\mathbf{\Sigma}_0\mathbf{\Lambda}^{-1}, \\ \mathbf{R}^d &= \mathbb{E}[(\mathbf{p} - \mathbf{p}_0)\mathbf{y}^\top] = \mathbf{\Lambda}\mathbb{E}[\mathbf{y}\mathbf{y}^\top] = 2\mathbf{\Lambda}\mathbf{\Omega} = \frac{1}{2}\mathbf{\Sigma}_0\mathbf{\Lambda}^{-1}, \\ \mathbf{R} &= \mathbb{E}[(\mathbf{p} - \mathbf{p}_0)\mathbf{x}^\top] = \mathbf{\Lambda}\mathbb{E}[\mathbf{x}\mathbf{x}^\top] = \mathbf{\Lambda}\mathbf{\Omega} = \frac{1}{4}\mathbf{\Sigma}_0\mathbf{\Lambda}^{-1}. \end{aligned}$$

□

An interesting consequence of this proposition is that within the present framework, the dressed responses should be entirely determined by $\mathbf{\Sigma}_0$ and $\mathbf{\Omega}$ alone. In Section 5 we will see how this naturally leads to the calibration of an impact model in which part of the estimation noise can be reduced by imposing the response structure above as a reasonable prior. Finally, note that while the dressed and bare responses of the fundamental price coincide, the dressed response of the traded price is a factor 2 larger than the bare response, due to the spurious correlations between prices and volume transiently introduced by the NT.

3.3 Utilities and competitive market making

In this model the utilities of the three agents are such that they sum to zero:

$$\begin{aligned} \mathcal{U}_{IT} &= -\mathbf{x}^\top(\mathbf{p} - \mathbf{v}), \\ \mathcal{U}_{NT} &= -\mathbf{u}^\top(\mathbf{p} - \mathbf{v}), \\ \mathcal{U}_{MM} &= (\mathbf{x} + \mathbf{u})^\top(\mathbf{p} - \mathbf{v}). \end{aligned}$$

It is thus interesting to characterise how utilities are transferred from one agent to the other at equilibrium.

Proposition 7 (Utilities at equilibrium). *The utilities of the three agents at equilibrium are equal to*

$$\mathbb{E}[\mathcal{U}_{IT}] = \text{tr}(\mathbf{R}), \quad (29)$$

$$\mathbb{E}[\mathcal{U}_{NT}] = -\text{tr}(\mathbf{R}), \quad (30)$$

$$\mathbb{E}[\mathcal{U}_{MM}] = 0. \quad (31)$$

Proof.

$$\begin{aligned}\mathbb{E}[\mathcal{U}_{IT}] &= -\mathbb{E}[\mathbf{x}^\top(\mathbf{p} - \mathbf{v})] = \text{tr}(\mathbb{E}[\mathbf{v}\mathbf{x}^\top] - \mathbb{E}[\mathbf{p}\mathbf{x}^\top]) = \text{tr}(\mathbf{R}_v - \mathbf{R}) = \text{tr}(\mathbf{R}), \\ \mathbb{E}[\mathcal{U}_{NT}] &= -\mathbb{E}[\mathbf{u}^\top(\mathbf{p} - \mathbf{v})] = \text{tr}(\mathbb{E}[\mathbf{v}\mathbf{u}^\top] - \mathbb{E}[\mathbf{p}\mathbf{u}^\top]) = \text{tr}(0 - \mathbf{R}) = -\text{tr}(\mathbf{R}), \\ \mathbb{E}[\mathcal{U}_{MM}] &= -\mathbb{E}[\mathcal{U}_{IT}] - \mathbb{E}[\mathcal{U}_{NT}] = 0.\end{aligned}$$

□

As in the standard Kyle model, the linear equilibrium of the Multivariate Kyle Model is such that wealth is transferred from the NT to the IT, which is able to capitalise his informational advantage. The role of the MM is more subtle: we assumed that the MM enforces an efficient price, hence by construction he is not optimising his wealth. However, as we show in Proposition 7, imposing price efficiency through the pricing rule (2) results in a break-even for the MM in the sense that the expectation of his utility is 0. In the next example we show that the contrary is not true: imposing the break-even condition for the MM is in general not sufficient to ensure efficient prices.

Example 3 (Efficient prices and break-even condition). *Consider the break-even condition $\mathbb{E}[\mathcal{U}_{MM}] = 0$. The expected utility of the MM in the univariate model reads:*

$$\mathbb{E}[\mathcal{U}_{MM}] = \frac{1}{2}(\mu - p_0)^2 \lambda^{-1} + \lambda \omega^d - r_v^d = 0.$$

Imposing $\mu = p_0$ rules out the presence of complex solutions regardless of r_v^d and ω^d and therefore:

$$\begin{aligned}\lambda &= r_v^d / \omega^d, \\ \mu &= p_0.\end{aligned}$$

Hence in one dimension a break-even condition for the MM, together with the requirement $\mu = p_0$ is equivalent to imposing price efficiency. This equivalence stems from the low dimensionality of the system because when we impose the break-even condition in one dimension we obtain one equation for one parameter (μ is fixed by imposing that the solution has to be real for any choice of parameters).

In the multivariate Kyle model this is not the case. In particular, the break-even condition writes

$$\mathbb{E}[\mathcal{U}_{MM}] = \text{tr} \left(\frac{1}{2}(\boldsymbol{\mu} - \mathbf{p}_0)(\boldsymbol{\mu} - \mathbf{p}_0)^\top \boldsymbol{\Lambda}_S^{-1} + \boldsymbol{\Lambda} \boldsymbol{\Omega}^d - \mathbf{R}_v \right) = 0.$$

In this case we obtain again one equation but we now need to fix $\sim n^2$ parameters. This means that there are many ways to break-even while only one is efficient. In a more restrictive setting where the MM has to break-even for each asset individually, the condition can be written as:

$$\text{diag} \left(\frac{1}{2}(\boldsymbol{\mu} - \mathbf{p}_0)(\boldsymbol{\mu} - \mathbf{p}_0)^\top \boldsymbol{\Lambda}_S^{-1} + \boldsymbol{\Lambda} \boldsymbol{\Omega}^d - \mathbf{R}_v \right) = 0. \quad (32)$$

In this case there are n equations, which are still not enough to fix $\sim n^2$ parameters.

Therefore, in one dimension the condition $\mathbb{E}[\mathcal{U}_{MM}] = 0$ is both a necessary and sufficient condition for the prices to be efficient. In multiple dimensions, although $\mathbb{E}[\mathcal{U}_{MM}] = 0$ is still a consequence of price efficiency, the converse is not true, as clarified in the next remark.

Remark 1 (Sufficient and necessary conditions for efficiency). *The asset-wise break even condition (32) can be also written as:*

$$\mathbb{E}[(p_i - v_i)y_i] = 0, \quad (33)$$

for each asset i . Under the assumption $\boldsymbol{\mu} = \mathbf{p}_0$ and in matrix form, (33) leads to

$$\begin{aligned} 0 &= \text{diag}(\mathbb{E}[(\mathbf{p} - \mathbf{v})\mathbf{y}^\top]) \\ &= \text{diag}(\boldsymbol{\Lambda}\mathbb{E}[\mathbf{y}\mathbf{y}^\top] - \mathbb{E}[\mathbf{v}\mathbf{y}^\top]) \\ &= \text{diag}(\boldsymbol{\Lambda}\boldsymbol{\Omega}^d - \mathbf{R}_v) \end{aligned}$$

which corresponds to the diagonal of (15). In order to recover (15) (and not only its diagonal) we need the off-diagonal version of equation (33):

$$\mathbb{E}[(p_i - v_i)y_j] = 0 ,$$

for $j \neq i$. Unfortunately, these terms are not directly related to any utility, and merely indicate that in order to prices to be efficient, the MM should destroy any correlation between the order flow y_j and the mispricing $p_i - v_i$.

Hence, whereas in one dimension the behavior of a Kyle MM can be justified on the basis of utility and competitive market making arguments (see Ref. [Kyl85]), it is much less clear why in the multi-asset case a competitive MM would enforce price efficiency on the basis of his/her own utility alone.

4 Interpretation

The properties of the linear equilibrium that have been illustrated in the section above provide some intuition about the phenomenology of the model, but they do not help in making sense of Eq. (17) for the impact $\boldsymbol{\Lambda}$. The following discussions are meant to provide a justification of the expression for $\boldsymbol{\Lambda}$, so to shed some light on its structure. As far as we know, the following arguments have never appeared in the literature before.

4.1 Whitening of prices and volumes

The derivation of the linear equilibrium presented in Theorem 1 relies on the quadratic matrix equation (16) which is solved by (see proof of Proposition 11):

$$\boldsymbol{\Lambda} = \frac{1}{2}\mathcal{G}\mathcal{O}\mathcal{L}^{-1}, \quad (34)$$

where we have used a decomposition for the price correlations $\boldsymbol{\Sigma}_0 = \mathcal{G}\mathcal{D}$ analogous to that of $\boldsymbol{\Omega} = \mathcal{L}\mathcal{R}$ and \mathcal{O} is an orthogonal matrix. The remarkable part of this finding is that, regardless of $\boldsymbol{\Sigma}_0$ and $\boldsymbol{\Omega}$, there always exists a unique change of basis \mathcal{O} such that $\boldsymbol{\Lambda}$ is SPD. Hence, the prediction process operated by the MM that results into the pricing rule $\Delta\mathbf{v} = \mathbf{v} - \mathbf{p}_0 = \boldsymbol{\Lambda}\mathbf{y}$ can be seen as resulting from the following steps:

1. Apply the *whitening* transformation \mathcal{L}^{-1} to \mathbf{y} , so to obtain the whitened imbalance $\tilde{\mathbf{y}} = 2^{-1/2}\mathcal{L}^{-1}\mathbf{y}$, with $\mathbb{C}[\tilde{\mathbf{y}}, \tilde{\mathbf{y}}] = \mathbb{I}$.
2. Apply the rotation \mathcal{O} to $\tilde{\mathbf{y}}$, obtaining the whitened fundamental price variations $\Delta\tilde{\mathbf{v}} = \mathcal{O}\tilde{\mathbf{y}}$. As before, one has in fact $\mathbb{C}[\Delta\tilde{\mathbf{v}}, \Delta\tilde{\mathbf{v}}] = \mathbb{E}[\Delta\tilde{\mathbf{v}}\Delta\tilde{\mathbf{v}}^\top] = \mathbb{I}$.
3. Apply the *inverse whitening* transformation \mathcal{G} in order to find the prediction for the fundamental price variation $\Delta\mathbf{v} = \mathbf{v} - \mathbf{p}_0 = 2^{-1/2}\mathcal{G}\Delta\tilde{\mathbf{v}}$.

While the first and the third steps of this procedure are intuitive, and can be seen as arising from dimensional analysis only, the nature of the rotation \mathcal{O} applied during the second step is less trivial, and will be investigated more closely in the following sections.

Example 4 (Degeneracy in the one-dimensional Kyle model). *In one dimension Eq. (16) becomes*

$$\frac{\sigma_0}{4} = \lambda^2\omega ,$$

leading to the couple of solutions $\lambda = \pm \frac{1}{2} \sqrt{\sigma_0/\omega}$. The degeneracy between the solutions is easily resolved by imposing positive-definiteness of the impact $\mathbf{\Lambda}$, which selects the positive solution. In this case we have $\mathcal{G} = \sqrt{\sigma_0}$, $\mathcal{L} = \sqrt{\omega}$, $\mathbf{O} = \mathbf{1}$. Hence, the one-dimensional Kyle model has too low a dimensionality for anything non-trivial to happen from the point of view of \mathbf{O} : it is only in higher dimension that one can appreciate its general structure.

4.2 Basis of prices and basis of volumes

As we show in the proof of Proposition 11, the rotation \mathbf{O} that appears in Eq. (34) can be expressed as

$$\mathbf{O} = (\mathcal{R}\mathcal{G})^{-1} \sqrt{(\mathcal{R}\mathcal{G})(\mathcal{R}\mathcal{G})^\top}. \quad (35)$$

In order to make some progress in understanding the complex nature of this rotation for $n > 1$, it is important to notice that it is completely specified in terms of the matrix $\mathcal{R}\mathcal{G}$. Let us assume that the matrix factorization chosen to compute $\mathbf{\Omega}$ and $\mathbf{\Sigma}_0$ is the Principal Component Decomposition:

$$\mathbf{\Omega} = \underbrace{\mathbf{W} \text{diag}(\sqrt{\omega})}_{\mathcal{L}} \underbrace{\text{diag}(\sqrt{\omega}) \mathbf{W}^\top}_{\mathcal{R}}, \quad (36)$$

$$\mathbf{\Sigma}_0 = \underbrace{\mathbf{S} \text{diag}(\sqrt{\sigma_0})}_{\mathcal{G}} \underbrace{\text{diag}(\sqrt{\sigma_0}) \mathbf{S}^\top}_{\mathcal{D}}, \quad (37)$$

where \mathbf{W} and \mathbf{S} are orthogonal matrices, and where the vectors $\sqrt{\omega}$ and $\sqrt{\sigma_0}$ have positive elements. Then we have:

$$\mathcal{R}\mathcal{G} = \text{diag}(\sqrt{\omega}) \mathbf{W}^\top \mathbf{S} \text{diag}(\sqrt{\sigma_0}), \quad (38)$$

indicating that, besides the diagonal matrices $\text{diag}(\sqrt{\omega})$ and $\text{diag}(\sqrt{\sigma_0})$ that set the scale of the fluctuations of fundamental prices and order imbalances, it is the overlap $\mathbf{W}^\top \mathbf{S}$ that links the eigenvectors of the volumes with that of the fundamental price. In order to obtain a stronger insight about the structure of \mathbf{O} , we shall proceed with some simple examples.

Example 5 (Two-dimensional Kyle model). *In two dimensions the matrix \mathbf{O} can be characterised, without loss of generality, by a single angle θ :*²

$$\mathbf{O} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}.$$

Consider a system given by the following correlations:

$$\mathbf{\Sigma}_0 = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}, \quad \mathbf{\Omega} = \begin{pmatrix} \omega_1 & 0 \\ 0 & \omega_2 \end{pmatrix}.$$

In this case we can calculate \mathbf{O} explicitly. In particular one has

$$\mathcal{G} = \frac{1}{\sqrt{2}} \begin{pmatrix} \sqrt{1+\rho} & -\sqrt{1-\rho} \\ \sqrt{1+\rho} & \sqrt{1-\rho} \end{pmatrix}, \quad \mathcal{L}^{-1} = \begin{pmatrix} \omega_1^{-1/2} & 0 \\ 0 & \omega_2^{-1/2} \end{pmatrix},$$

and imposing the symmetry of $\mathbf{\Lambda}$ one can show that

$$\theta = \arcsin \left(\frac{\omega_1^{-1/2} \sqrt{1+\rho} + \omega_2^{-1/2} \sqrt{1-\rho}}{\sqrt{2}\Delta} \right)$$

where $\Delta = \sqrt{\omega_1 + \omega_2 + 2(\omega_1\omega_2)^{1/2}\sqrt{1-\rho^2}}$. Finally we obtain

$$\mathbf{\Lambda} = \frac{1}{2\Delta} \begin{pmatrix} 1 + \sqrt{\frac{\omega_2}{\omega_1}} \sqrt{1-\rho^2} & \rho \\ \rho & 1 + \sqrt{\frac{\omega_1}{\omega_2}} \sqrt{1-\rho^2} \end{pmatrix}. \quad (39)$$

²Indeed, by consistently choosing the factorisation of \mathcal{G} and \mathcal{L} one can fix arbitrarily the determinant of \mathbf{O} to be plus or minus one due to Eq. (35).

We can use the results in example 5 to examine a couple of extreme cases: the one of extreme illiquidity and the one in which assets are strong correlations.

Example 6 (Extreme illiquidity). *Let us consider the limit in which $\omega_2 = \epsilon\omega$ with $\epsilon \rightarrow 0$, whereas $\omega_1 = \omega$, implying that $y_2\epsilon^{-1/2}$ is expected to be finite. Then the prediction of the MM up to first order in ϵ is:*

$$\mathbf{p} - \mathbf{p}_0 = \frac{1}{2\sqrt{\omega}} \left(\rho y_1 + \sqrt{1 - \rho^2} (y_2 \epsilon^{-1/2}) \right),$$

implying that the efficient price of a liquid instrument is fixed solely by the volume traded on that same instrument, whereas for illiquid markets it is important to take into account quantities traded on liquid, correlated markets. On the other hand, the behavior of the IT up to first order in ϵ is given by:

$$\mathbf{x} = \sqrt{\omega} \left(\begin{array}{c} \Delta v_1 \\ \sqrt{\frac{\epsilon}{1-\rho^2}} (\Delta v_2 - \rho \Delta v_1) \end{array} \right).$$

Now the IT is encouraged to trade only the liquid asset and furthermore ignore the illiquid one when placing the bid. Similarly to the previous case, the traded price will remain of order one because the market orders for the second asset will be order $\sqrt{\epsilon}$.

Price innovation in example 6 can be interpreted as an indication of how an efficient market should work: the symmetry of cross-impact implies that the effect of trading one dollar of a very liquid asset a_1 on the price of an illiquid asset a_2 is the same as the effect of trading a dollar of asset a_2 on the price of asset a_1 , however because there are going to be very few dollars traded on a_2 it will not affect significantly the price of a_1 . Conversely, the price of a_2 will be heavily driven by the traded volume of a_1 .

Although we only proved the existence of the linear equilibrium for invertible Σ_0 , we can study the behavior of $\mathbf{\Lambda}$ in the limiting case in which Σ_0 has low rank as in the two following examples:

Example 7 (Strongly-correlated prices). *Consider $\rho = 1 - \delta$, with $\delta \rightarrow 0$, so that $(\Delta p_1 - \Delta p_2)/\sqrt{\delta}$ is finite. According to (39) the prediction of the MM up to first order in δ is:*

$$\mathbf{p} - \mathbf{p}_0 = \frac{1}{2\sqrt{\omega_1 + \omega_2}} \begin{pmatrix} y_1 + y_2 \\ y_1 + y_2 \end{pmatrix},$$

implying that when dealing with strongly correlated instruments, the efficient prices can be built by summing algebraically the volumes traded on each of them before normalizing by the global liquidity. On the other hand, the bet of the IT up to first order in δ can be written as:

$$\mathbf{x} = \frac{1}{2} \sqrt{\frac{1}{\omega_1 + \omega_2}} \begin{pmatrix} \omega_1 (\Delta v_1 + \Delta v_2) \\ \omega_2 (\Delta v_1 + \Delta v_2) \end{pmatrix} + \sqrt{\frac{\omega_1 \omega_2}{2(\omega_1 + \omega_2)}} \begin{pmatrix} \frac{1}{\sqrt{\delta}} (\Delta v_1 - \Delta v_2) \\ \frac{1}{\sqrt{\delta}} (\Delta v_2 - \Delta v_1) \end{pmatrix}.$$

In this case the IT is encouraged to place orders proportional to either the average price variation, or orders of finite size proportional to the relative difference, rescaled by the typical size of the relative price moves $\sqrt{\delta}$. Indeed, the IT should bet on price variations inversely proportional to how common they are.

The specialization of the result above to the extreme case $\rho \rightarrow \pm 1$ regime is particularly interesting, because it is a consequence of a more general result that can be applied whenever Σ_0 is rank one, which is shown in the next example.

Example 8 (Rank one price covariance). *Let us consider the case in which all eigenvalues of Σ_0 except for one tend to 0, and therefore Σ_0 tends to a rank 1 matrix. The interest of this example resides in the fact that $\mathbf{\Lambda}$ displays in this case a particularly illuminating expression, providing a simple recipe for pooling together volumes belonging to different financial instruments correlated to the same underlying product. Without loss of*

generality, Σ_0 in the rank one case can be written as $\Sigma_0 = \sigma \mathbf{s} \mathbf{s}^\top$ where $\mathbf{s} \in \mathbb{R}^n$ is a unit vector and $\sigma > 0$. It is then straightforward to verify that the matrix

$$\mathbf{\Lambda} = \frac{1}{2} \mathbf{s} \left(\frac{\sigma}{\mathbf{s}^\top \mathbf{\Omega} \mathbf{s}} \right)^{1/2} \mathbf{s}^\top \quad (40)$$

yields the linear equilibrium of the model. The above equation has a very simple interpretation:

- The matrix $\mathbf{\Lambda}$ commutes with Σ_0 , and its eigenvectors are insensitive to those of $\mathbf{\Omega}$ (i.e., the volume fluctuations induced by noise traders).
- The factor $\sigma^{1/2}$ sets the scale for the price variations as being that of the only mode of Σ_0 that has non-zero fluctuations.
- By writing the principal component decomposition of $\mathbf{\Omega}$ as $\mathbf{\Omega} = \sum_{a=1}^n \mathbf{w}_a \omega_a \mathbf{w}_a^\top$, one can write the denominator as:

$$(\mathbf{s}^\top \mathbf{\Omega} \mathbf{s})^{1/2} = \left(\sum_{a=1}^n (\mathbf{w}_a^\top \mathbf{s})^2 \omega_a \right)^{1/2}. \quad (41)$$

Such a denominator sets the scaling with volume of the impact $\mathbf{\Lambda}$. It amounts to a projection of $\mathbf{\Omega}$ on the only non-zero mode of Σ_0 , or equivalently a sum of the individual volume variances ω_a of $\mathbf{\Omega}$, weighted by their projections on \mathbf{s} .

5 Implications

The expression (17) for the impact matrix $\mathbf{\Lambda}$ is of interest beyond the context of the Kyle Model. It provides an insightful inference prescription for cross-impact models, as opposed to what has been done so far in the context of cross-impact fitting (see [BMEB17, MBEB17, WSG16, SL18]). In such a context, one is interested in estimating from empirical data a model to forecast a price variation $\Delta \mathbf{p}$ with a predictor of the form:

$$\hat{\Delta} \mathbf{p} = \hat{\mathbf{\Lambda}} \mathbf{y}, \quad (42)$$

in which it is essential to faithfully model how trading the instrument i will impact the instrument j . The goal of this section is to establish a link between the Kyle model and the empirical calibration of Eq. (42), where the price variations $\Delta \mathbf{p}$ and the imbalances \mathbf{y} are sampled from empirical data.

5.1 From theory to data: empirical averages and loss function

First, let us consider a dataset of T i.i.d. efficient price variations and volumes, $\{\Delta \mathbf{p}^{(t)}\}_{t=1}^T$ and $\{\mathbf{y}^{(t)}\}_{t=1}^T$, sampled from an unknown distribution with bounded variance that needs not to be related to the Kyle model presented in the previous sections of the manuscript. Accordingly, with a slight abuse of notation, from now on $\mathbb{E}[\cdot]$ denotes the averages taken with respect to the underlying distribution from which $\Delta \mathbf{p}$ and \mathbf{y} are sampled. In addition, we introduce an empirical measure $\langle \cdot \rangle$ to denote the averages with respect to the empirical sample under investigation. In particular, simple empirical estimators for the average price variation and the mean imbalance write:

$$\hat{\mathbf{p}}_0 = \langle \Delta \mathbf{p} \rangle = \frac{1}{T} \sum_{t=1}^T \Delta \mathbf{p}^{(t)} \quad (43)$$

$$\hat{\mathbf{y}}_0 = \langle \mathbf{y} \rangle = \frac{1}{T} \sum_{t=1}^T \mathbf{y}^{(t)}. \quad (44)$$

Henceforth, we consider that price changes and order flows are shifted by their empirical mean $\hat{\boldsymbol{p}}_0$ and $\hat{\boldsymbol{y}}_0$, and therefore we set $\hat{\boldsymbol{p}}_0 = \hat{\boldsymbol{y}}_0 = 0$. The corresponding covariances are then defined accordingly as:

$$\hat{\boldsymbol{\Sigma}} = \langle \Delta \boldsymbol{p} \Delta \boldsymbol{p}^\top \rangle, \quad (45)$$

$$\hat{\boldsymbol{\Omega}}^d = \langle \boldsymbol{y} \boldsymbol{y}^\top \rangle, \quad (46)$$

$$\hat{\boldsymbol{R}}^d = \langle \Delta \boldsymbol{p} \boldsymbol{y}^\top \rangle. \quad (47)$$

As $T \rightarrow \infty$, the empirical averages converge to the actual means and covariances. Our goal is to show different recipes for the calibration of the model (42), all relying on the estimators above. We shall compare such estimators with the *Kyle estimator* $\hat{\boldsymbol{\Lambda}}_{\text{Kyle}}$, which, as we shall see, displays several interesting properties. In order to evaluate the quality of different calibrations, we consider a quadratic loss χ^2 defined below.

Definition 2 (Loss). *Given an empirical measure $\langle \cdot \rangle$, we define the loss χ^2 as the function*

$$\chi^2 = \frac{1}{2} \left\langle (\hat{\Delta \boldsymbol{p}} - \Delta \boldsymbol{p})^\top \boldsymbol{M} (\hat{\Delta \boldsymbol{p}} - \Delta \boldsymbol{p}) \right\rangle, \quad (48)$$

where $\hat{\Delta \boldsymbol{p}} = \hat{\boldsymbol{\Lambda}} \boldsymbol{y}$ is a linear predictor of the efficient price variation and \boldsymbol{M} is a SPD matrix.

This definition of the loss implicitly implies that the calibration of the model is very similar to the task of the market maker in the Kyle model: whereas in the latter the job of the MM is to forecast what the *fundamental* price variation is on the basis of the order flow imbalance, in this setting one is required to forecast the *efficient* (observed) price on the basis of the imbalance. Even though the two problems are different, they involve extremely similar equations, a fact that will allow us to leverage the results of the previous sections, valid in principle for a Kyle Market Maker, in the calibration setting. Hence, under this definition of loss one is only required to find proxies for the price variations $\Delta \boldsymbol{p}$ and the dressed order imbalance \boldsymbol{y} , disregarding in principle the realisations of the underlying fundamental price \boldsymbol{v} .

5.2 Cross-impact estimators

Here we present three different cross-impact estimators and discuss their properties. The proofs of the propositions in this section are provided in App. B.

5.2.1 Maximum Likelihood Estimation

The simplest estimation recipe that one can construct in order to estimate the cross-impact matrix $\boldsymbol{\Lambda}$ is probably the Maximum Likelihood Estimator (MLE): $\hat{\boldsymbol{\Lambda}}_{\text{MLE}}$, that is obtained by minimising a quadratic loss.

Proposition 8. *The minimisation of the loss χ^2 with respect to $\hat{\boldsymbol{\Lambda}}_{\text{MLE}}$ yields:*

$$\hat{\boldsymbol{\Lambda}}_{\text{MLE}} = \hat{\boldsymbol{R}}^d (\hat{\boldsymbol{\Omega}}^d)^{-1}, \quad (49)$$

independent of \boldsymbol{M} . The loss at the minimum is given by:

$$\chi_{\text{MLE}}^2 = \frac{1}{2} \text{tr} \left(\boldsymbol{M} (\hat{\boldsymbol{\Sigma}} - \hat{\boldsymbol{\Sigma}}_{\text{MLE}}) \right), \quad (50)$$

where $\hat{\boldsymbol{\Sigma}}_{\text{MLE}} = \hat{\boldsymbol{R}}^d (\hat{\boldsymbol{\Omega}}^d)^{-1} (\hat{\boldsymbol{R}}^d)^\top$ is the portion of the covariance of fundamental price variations explained by the model (49).

Note that Eq. (49) is almost identical to Eq. (15), that has been obtained by enforcing the efficient price condition. This is no coincidence, as in a linear equilibrium setting with normally distributed variables the traded price is exactly a linear function of the order flow imbalance \boldsymbol{y} . There are though two important differences distinguishing the linear equilibrium of the Kyle model and the minimization of χ_{MLE}^2 :

- A MM in the Kyle setting is actually performing a linear regression of the *fundamental* price, whereas here we are interested in regressing the *efficient* price. This justifies why Eq. (15) uses the *fundamental* price response \mathbf{R}_v^d , as opposed to Eq. (49) that uses the efficient price \mathbf{R}^d .
- In the Kyle setup case, the MM is aware that a part of the order imbalance comes from the IT who owns information about the fundamental price and is optimising his strategy assuming that the MM uses a MLE. This leads to a quadratic equation for $\mathbf{\Lambda}$ (see Eq. (57) below) even in the linear equilibrium setting, whereas Eq. (49) above is a linear relation for $\hat{\mathbf{\Lambda}}_{\text{MLE}}$.

Remark 2. *One could have defined the Multivariate Kyle Model by replacing the efficient-price condition (see Eq. (2)) with a condition on the minimisation of χ^2 without affecting the linear equilibrium. Moreover, this approach would do a better job at generalising the model to more exotic settings. In fact, Eq. (??) provides a linear predictor only under the assumption of normality for prices and volumes: for other distributions, the relation linking \mathbf{y} and \mathbf{v} is in general non-linear. Having a MM that minimises a loss χ^2 is a reasonable assumption if one thinks that a MM without enough knowledge of the underlying distributions or computational power to build an efficient price should rely on linear regressions in order to estimate the latter.*

The expression for $\hat{\mathbf{\Lambda}}_{\text{MLE}}$ only contains the dressed empirical estimators $\hat{\mathbf{R}}^d$ and $\hat{\mathbf{\Omega}}^d$, allowing one to estimate the impact matrix from real data. In addition, such an estimator has the benefit of being very simple to implement, as it only requires the solution of a linear equation for $\hat{\mathbf{\Lambda}}_{\text{MLE}}$. Unfortunately, this estimator lacks several properties that turn out to be very useful in cases of practical interest:

Symmetry The matrix $\hat{\mathbf{\Lambda}}_{\text{MLE}}$ is symmetric if and only if $\hat{\mathbf{R}}^d$ and $\hat{\mathbf{\Omega}}^d$ commute.

Positive definiteness The matrix $\hat{\mathbf{\Lambda}}_{\text{MLE}}$ can have negative eigenvalues (see [BMEB17]).

Consistency of correlations In general, $[\hat{\mathbf{\Sigma}}, \hat{\mathbf{\Sigma}}_{\text{MLE}}] \neq 0$. Hence, the price variations inferred by using the order flow imbalance do not share the eigenvectors of the real price variations, unless the response $\hat{\mathbf{R}}^d(\hat{\mathbf{\Omega}}^d)^{-1}\hat{\mathbf{R}}^d$ commutes with $\hat{\mathbf{\Sigma}}$.

The first two properties are extremely important in the calibration of cross-impact models, as shown in [AKS16, SL18]: to perform pricing within a cross-impact setup, the matrix $\mathbf{\Lambda}$ should be SPD in order to ensure absence of price manipulation. This is not the case for a MLE, which is thus not suitable for practical purposes.

5.2.2 EigenLiquidity model

In order to cure the lack of symmetry of the Maximum Likelihood Estimator, the idea proposed in Ref. [MBEB17] is to construct an estimator of cross-impact $\hat{\mathbf{\Lambda}}_{\text{ELM}}$ that is symmetric by construction, by enforcing the relation:

$$[\hat{\mathbf{\Lambda}}_{\text{ELM}}, \hat{\mathbf{\Sigma}}] = 0, \quad (51)$$

which means that the impact eigendirections coincide with the eigenportfolios (or principal components of the asset space). This prevents price manipulations that would be induced by an asymmetric of $\mathbf{\Lambda}$. Its calibration is explained in the next proposition.

Proposition 9. *Consider a pricing rule $\hat{\Delta}\mathbf{p} = \hat{\mathbf{\Lambda}}_{\text{ELM}}\mathbf{y}$, in which we impose the commutation relation (51). Then, the maximum likelihood estimator obtained under such constraint takes the form:*

$$\hat{\mathbf{\Lambda}}_{\text{ELM}} = \sum_{a=1}^n \hat{\mathbf{s}}_a g_a \hat{\mathbf{s}}_a^\top, \quad (52)$$

where $\{\hat{\mathbf{s}}_a\}_{a=1}^n$ are the empirical eigenvectors of $\hat{\mathbf{\Sigma}}$, and where:

$$g_a = \frac{\hat{\mathbf{s}}_a^\top \hat{\mathbf{R}}^d \hat{\mathbf{s}}_a}{\hat{\mathbf{s}}_a^\top \hat{\mathbf{\Omega}}^d \hat{\mathbf{s}}_a}. \quad (53)$$

Furthermore, the loss at the minimum is given by:

$$\chi_{\text{ELM}}^2 = \frac{1}{2} \text{tr} \left[M \left(\hat{\Sigma} - \hat{\mathbf{R}}^{\text{d}} \sum_{a=1}^n \hat{\mathbf{s}}_a \frac{\hat{\mathbf{s}}_a^\top \hat{\mathbf{R}}^{\text{d}} \hat{\mathbf{s}}_a}{\hat{\mathbf{s}}_a^\top \hat{\Omega}^{\text{d}} \hat{\mathbf{s}}_a} \hat{\mathbf{s}}_a^\top \right) \right]. \quad (54)$$

In this case, the properties discussed above become:

Symmetry The matrix $\hat{\Lambda}_{\text{ELM}}$ is symmetric by construction.

Positive definiteness The matrix $\hat{\Lambda}_{\text{ELM}}$ can still have negative eigenvalues, although empirically the matrix $\hat{\Lambda}_{\text{ELM}}$ has been reported not to display eigenvalues significantly smaller than zero (see [MBEB17]).

Consistency of correlations Similarly to the MLE case, the price variation covariances $\hat{\Sigma}_{\text{ELM}} = \hat{\Lambda}_{\text{ELM}} \hat{\Omega}^{\text{d}} \hat{\Lambda}_{\text{ELM}}$ do not generally commute with $\hat{\Sigma}$. It is the case if and only if $[\hat{\Omega}^{\text{d}}, \hat{\Sigma}] = 0$.

Summarizing, the price to pay in order to have a symmetric estimator is a larger loss function. Note that if $\hat{\mathbf{R}}^{\text{d}}$ and $\hat{\Omega}^{\text{d}}$ commute with $\hat{\Sigma}$, then $\hat{\Lambda}_{\text{ELM}} = (\hat{\mathbf{R}}^{\text{d}})^2 (\hat{\Omega}^{\text{d}})^{-1} = \hat{\Lambda}_{\text{MLE}}$ and therefore the loss is equal to the best possible $\chi_{\text{ELM}}^2 = \chi_{\text{MLE}}^2$.

Remark 3 (Estimators in one dimension). *Since commutation and symmetry are granted in one dimension, imposing the condition (51) does not add any constraints to the minimization of the loss function and therefore $\hat{\Lambda}_{\text{ELM}} = \hat{\Lambda}_{\text{MLE}} = \frac{r^{\text{d}}}{\omega^{\text{d}}}$.*

5.2.3 Kyle estimator

Now, let us provide some intuition on how to construct an impact estimator inspired by Kyle's model, $\hat{\Lambda}_{\text{Kyle}}$, by characterising the Λ in the Kyle model from a different perspective. First, take the MM solution Eq. (15) (matching the MLE, Eq. (49)), and exchange the dressed order imbalances and responses with those predicted in the linear equilibrium through Eqs. (28) and (23):

$$\mathbf{R}_v^{\text{d}} \rightarrow \frac{1}{2} \Sigma_0 \Lambda^{-1}, \quad (55)$$

$$\Omega^{\text{d}} \rightarrow 2\Omega. \quad (56)$$

Then, the linear regression of Eq. (15) becomes a quadratic matrix equation:

$$\Lambda = \mathbf{R}_v^{\text{d}} (\Omega^{\text{d}})^{-1} \rightarrow \Lambda = \frac{1}{4} \Sigma_0 \Lambda^{-1} \Omega^{-1}. \quad (57)$$

After imposing symmetry and positive definiteness of Λ , this is exactly the equation that leads to the expression of Λ obtained in the Kyle linear equilibrium (see App. A.4). By doing so, one ensures an efficient price variation process whose associated covariance is half that of the ‘‘true’’ price variations. As alluded to above, in the Kyle setup the MM is implicitly performing a linear regression, and *knows that the IT is aware of his/her algorithm*, finally leading to Eq. (57). We now propose an estimator $\hat{\Lambda}_{\text{Kyle}}$ that exploits these ideas.

Proposition 10. *Let us consider an estimator $\hat{\Delta p} := \hat{\Lambda}_{\text{Kyle}} \mathbf{y}$ such that*

- $\hat{\Lambda}_{\text{Kyle}}$ is SPD
- The empirical covariance of the associated efficient price variations $\hat{\Sigma}_{\text{Kyle}} = \langle \hat{\Delta p} \hat{\Delta p}^\top \rangle$ satisfies

$$\hat{\Sigma}_{\text{Kyle}} = k^2 \hat{\Sigma}$$

with $k \in \mathbb{R}$. Then the unique $\hat{\Lambda}_{\text{Kyle}}$ satisfying these constraints is given by

$$\hat{\Lambda}_{\text{Kyle}} = k (\hat{\mathcal{R}}^{\text{d}})^{-1} \sqrt{\hat{\mathcal{R}}^{\text{d}} \hat{\Sigma} \hat{\mathcal{L}}^{\text{d}}} (\hat{\mathcal{L}}^{\text{d}})^{-1} \quad (58)$$

where $\hat{\Omega}^d = \hat{\mathcal{L}}^d \hat{\mathcal{R}}^d$ is a decomposition of $\hat{\Omega}^d$ such that $\hat{\mathcal{L}}^d = (\hat{\mathcal{R}}^d)^\top$.
Moreover, the loss at the minimum is

$$\chi_{\text{Kyle}}^2 = \frac{1}{2} \text{tr} \left[\mathbf{M} \left((1+k^2) \hat{\Sigma} - 2k \hat{\mathbf{R}}^d (\hat{\mathcal{R}}^d)^{-1} \sqrt{\hat{\mathcal{R}}^d \hat{\Sigma} \hat{\mathcal{L}}^d (\hat{\mathcal{L}}^d)^{-1}} \right) \right] \quad (59)$$

and the value of k that minimizes the loss is given by

$$k^* = \frac{\text{tr}(\mathbf{M} \hat{\mathbf{R}}^d (\hat{\mathcal{R}}^d)^{-1} \sqrt{\hat{\mathcal{R}}^d \hat{\Sigma} \hat{\mathcal{L}}^d (\hat{\mathcal{L}}^d)^{-1}})}{\text{tr}(\mathbf{M} \hat{\Sigma})}.$$

Note that the complete analogy with the Kyle model is recovered for $k = 1$. In fact, if the sample dataset is drawn from an actual Multivariate Kyle Model with $\Sigma_0/2 = \Sigma = \hat{\Sigma}$ and $2\Omega = \Omega^d = \hat{\Omega}^d$ as described in the previous sections, one obtains:

$$\mathbb{E}[\hat{\mathbf{\Lambda}}_{\text{Kyle}}] \xrightarrow{T \rightarrow \infty} k \mathbf{\Lambda}.$$

The advantage of leaving k as a free parameter is that it allows to increase or decrease the loss χ^2 of the Kyle estimator without affecting the eigenvectors of the predicted covariance. In fact, the relationship $2\Sigma = \Sigma_0$ arising in the Kyle model is thought not to be universal, so that it is more natural to think of it as a phenomenological parameter expressing the informational content of trades.

The adoption of $\hat{\mathbf{\Lambda}}_{\text{Kyle}}$ as an estimator of cross-impact has several interesting advantages with respect to more customary estimators such as the Maximum Likelihood Estimator $\hat{\mathbf{\Lambda}}_{\text{MLE}}$ or an impact estimator based on the Eigen-Liquidity Model $\hat{\mathbf{\Lambda}}_{\text{ELM}}$:

Symmetry and positive definiteness The Kyle construction imposes *a priori* symmetry and PDness due to the request of optimality of the IT. In fact, the optimality condition (14) and the stability one (PDness of $\mathbf{\Lambda}$) are exactly the ones that a risk-neutral rational agent, as considered in [AKS16, SL18], would face when trying to optimise his profits.

Consistency of correlations This construction allows to recover the empirical correlations $\hat{\Sigma}$ for $k = 1$. For any value of k one has $[\hat{\Sigma}, \hat{\Sigma}_{\text{Kyle}}] = 0$.

Loss An important point concerns the loss function obtained by using $\hat{\mathbf{\Lambda}}_{\text{Kyle}}$. In general χ_{Kyle}^2 is *larger* than that obtained by doing Maximum Likelihood Estimation, which minimizes χ^2 by design. Hence, the price to pay in order to have symmetry, positive semi-definiteness and consistency of correlations is in general a worse fit of empirical data with respect to $\hat{\mathbf{\Lambda}}_{\text{MLE}}$ (as measured by the loss χ^2). However, if our data behaves as the Kyle model, i.e. if $\hat{\mathbf{R}}^d = \hat{\Sigma} \hat{\mathbf{\Lambda}}_{\text{Kyle}}^{-1}$ then $k^* = 1$ and:

$$\chi_{\text{Kyle}}^2 = 0 = \chi_{\text{MLE}}^2,$$

implying that $\mathbf{\Lambda}$ is (*obviously!*) the best estimator for a market that reproduces the statistics of the multivariate Kyle model. Moreover, the loss at the minimum is zero because the *efficient* price variation $\Delta \mathbf{p}$ is completely determined by the imbalance. Indeed, the problem of regressing the *fundamental* price variation $\Delta \mathbf{v}$ would have yielded a non-zero loss at the minimum even for data sampled from an actual Kyle model due to the relation $\Sigma < \Sigma_0$.

Remark 4 (ELM and Kyle estimators for rank one $\hat{\Sigma}$). *As we showed in example 8, when the price variation covariance is rank one $\hat{\mathbf{\Lambda}}_{\text{Kyle}}$ is proportional to $\hat{\Sigma}$. On the other hand, $\hat{\mathbf{\Lambda}}_{\text{ELM}}$ has, by definition, the same eigenbasis as the price variations covariance. Therefore, if $\hat{\Sigma}$ is rank one then $\hat{\mathbf{\Lambda}}_{\text{ELM}}$ is proportional to $\hat{\mathbf{\Lambda}}_{\text{Kyle}}$ and $\hat{\Sigma}$.*

5.3 A comparison of recipes

In order to compare the performance of the estimators we compute observables that are directly related to the properties discussed above:

- **Loss:** we compute the loss χ^2 as given in Eq. (48).
- **Asymmetry:** to quantify the degree of symmetry of the estimator we compute the norm of its antisymmetric part divided by the total norm:

$$\alpha = \frac{|\hat{\mathbf{\Lambda}} - \hat{\mathbf{\Lambda}}^\top|}{2|\hat{\mathbf{\Lambda}}|}.$$

The value $\alpha = 0$ ($\alpha = 1$) means that the estimator is symmetric (antisymmetric).

- **Positive definiteness:** we compute the lowest real part of the spectrum of the estimator:

$$\lambda^* = \min_i \Re(\hat{\lambda}_i),$$

where $\{\hat{\lambda}_i\}_{1 \geq i \geq n}$ are the eigenvalues of $\hat{\mathbf{\Lambda}}$. Positive definiteness is equivalent to $\lambda^* > 0$.

- **Consistency of correlations:** we compute the norm of the commutator of $\hat{\mathbf{\Sigma}}$ and the covariances of the estimated price variations $\hat{\mathbf{\Sigma}}_{\text{est}}$:

$$\kappa = \frac{|\hat{\mathbf{\Sigma}}_{\text{est}} \hat{\mathbf{\Sigma}} - \hat{\mathbf{\Sigma}} \hat{\mathbf{\Sigma}}_{\text{est}}|}{|\hat{\mathbf{\Sigma}}| |\hat{\mathbf{\Sigma}}_{\text{est}}|}.$$

Correlations are consistent if $\kappa = 0$.

5.3.1 Application to real data

In Fig. 1 we show the covariances and estimators for the 2-year, 5-year, 10-year, and 30-year tenors of the U.S. Treasury Futures traded in the Chicago Board of Trade (see B.5 for details). In Table 1 we show the values of the observables corresponding to the 4×4 system. In Table 2 we show the values of the same observables but averaged over the 6 possible combinations of 2×2 systems. As expected, the Kyle procedure fares slightly worse for the loss function, but is much better than other procedures on all other counts.

Estimator	Loss (χ^2)	Commutator (κ)	PDness (λ^*)	Asymmetry (α)
$\hat{\mathbf{\Lambda}}_{\text{MLE}}$	1.20	0.038	0.031	0.31
$\hat{\mathbf{\Lambda}}_{\text{ELM}}$	1.26	0.129	0.017	0
$\hat{\mathbf{\Lambda}}_{\text{Kyle}}$	1.32	0	0.386	0

Table 1: Values of the observables for the three estimators for U.S. Treasury Futures and Bonds.

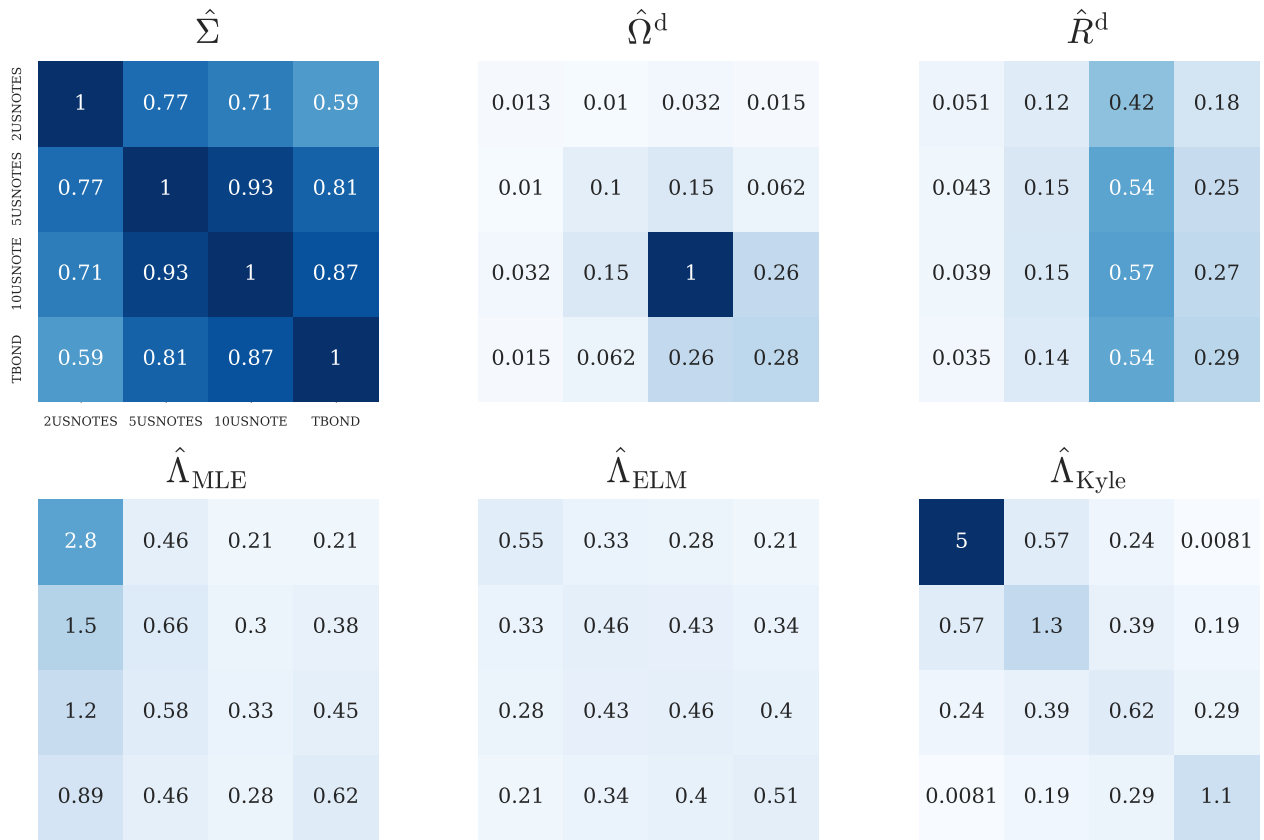


Figure 1: **Averaged US Treasury Futures covariances and impact estimators for 2016.**

Top from left to right: Daily averaged price variation covariances in units of risk squared. Unitless daily averaged volume covariances rescaled with respect to the maximum. To obtain the original volumes $\hat{\Omega}^d$ has to be multiplied by $2242\2 . Market response rescaled with respect to the maximal volume in units of risk. To obtain the original volumes \hat{R}^d has to be multiplied by $\sqrt{2242}\$$.

Bottom from left to right: Maximum Likelihood impact estimator, EigenLiquidity Model based impact estimator and Kyle model based impact estimator in units of risk. To obtain the estimators in the right units they have to be divided by $\sqrt{2242}\$$.

Estimator	Loss (χ^2)	Commutator (κ)	PDness (λ^*)	Asymmetry (α)
$\hat{\Lambda}_{MLE}$	0.668	0.033	0.318	0.186
$\hat{\Lambda}_{ELM}$	0.691	0.108	0.231	0
$\hat{\Lambda}_{Kyle}$	0.718	0	0.952	0

Table 2: Average values of the observables for the three estimators for the 6 pairs of U.S. Treasury Futures.

5.3.2 Synthetic data

To further explore the effects of price variation and liquidity correlations, in the next two sections we compute the values of the observables described above on the 6 sets of synthetic 2×2 covariance matrices with specific

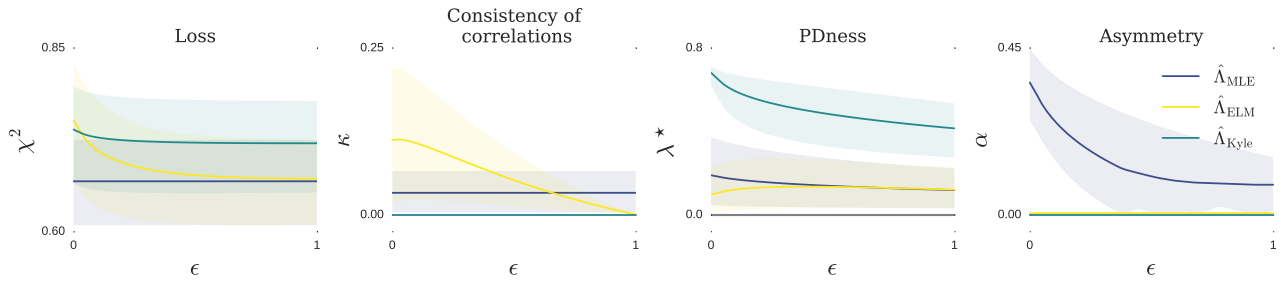


Figure 2: **Effect of liquidity.** Averaged values (solid lines) and intervals between the maximum and the minimum (shaded area) of the observables for ϵ ranging from 0 to 1.

price variation covariances ($\hat{\Sigma}_\rho, \hat{\Omega}_\rho^d, \hat{R}_\rho^d$) and specific volume covariances ($\hat{\Sigma}_\epsilon, \hat{\Omega}_\epsilon^d, \hat{R}_\epsilon^d$) respectively fabricated by modifying the data for U.S. Treasury Futures as explained in B.5 (see Table 2). The idea behind this construction is to provide a synthetic but realistic set of matrices that are parametrized by either a liquidity parameter ϵ or a correlation parameter ρ . By varying those parameters, one can extrapolate from the reality (recovered for specific values of ϵ and/or ρ) and a fictitious world in which assets can be made more or less liquid with respect to reality by varying the knob ϵ , and more or less correlated by varying the parameter ρ .

Extreme illiquidity In order to explore the effect of extreme heterogeneous liquidities we rescale the 2×2 covariances as:

$$\hat{\Omega}^d \mapsto \hat{\Omega}_\epsilon^d = \begin{pmatrix} 1 & \sqrt{\epsilon} \hat{\Omega}_{12}^d \\ \sqrt{\epsilon} \hat{\Omega}_{12}^d & \epsilon \end{pmatrix}.$$

In Fig. 2 we show the value of each of the observables averaged over the 6 sets of covariances for ϵ ranging from 0 to 1. Regarding the loss, the estimator that better performs is $\hat{\Lambda}_{MLE}$, regardless of ϵ as expected. It is interesting to observe that χ_{MLE}^2 does not depend on ϵ . The reason is that, by construction, the estimated price variation covariances are invariant under changes in the volume variances³:

$$\hat{\Sigma}_{MLE} = \hat{R}_\epsilon^d (\hat{\Omega}_\epsilon^d)^{-1} (\hat{R}_\epsilon^d)^\top = \hat{R}^d (\hat{\Omega}^d)^{-1} (\hat{R}^d)^\top.$$

The loss of the other two estimators depends on ϵ . For small values of ϵ , χ_{Kyle}^2 is lower than χ_{ELM}^2 , suggesting that $\hat{\Lambda}_{Kyle}$ might be better in case of markets with heterogeneous volatilities. For $\epsilon \gtrsim 0.2$ the loss distributions for the three estimators have a lot of overlap and no one is significantly better. As explained above, κ_{MLE} does not depend on ϵ and $\kappa_{Kyle} = 0$, by construction. On the other hand, the limit κ_{ELM} decreases to 0 as $\epsilon \rightarrow 1$ because in the limit $\epsilon = 1$, $\hat{\Omega}_\epsilon^d$ and $\hat{\Sigma}_\epsilon$ commute. All estimators are PD, and the only non-symmetric estimator is $\hat{\Lambda}_{MLE}$, whose antisymmetric part decreases as ϵ increases (even though it never becomes fully symmetric). The other two are symmetric by construction.

Reminiscent to what we showed in example 6, in Fig. 3 we display the covariance matrices and the three corresponding impact estimators in a case where one of the assets' liquidity is much smaller than the other one. Note that due to the lack of liquidity of the second asset the second column of \hat{R}_ϵ^d is considerably smaller than the first one while price covariances are not changed. Let us analyze heuristically the estimators obtained:

- $\hat{\Lambda}_{MLE}$: Trading the second asset will have a strong impact on its price and a significant impact on the price of first asset (due to the non-negligible price correlations). Trading the first asset will have little impact on both prices.
- $\hat{\Lambda}_{ELM}$: In this case the impact estimator predicts that trading either asset will have a similar impact.

³This is not a particular property of systems in 2 dimensions, it holds for all dimensions.

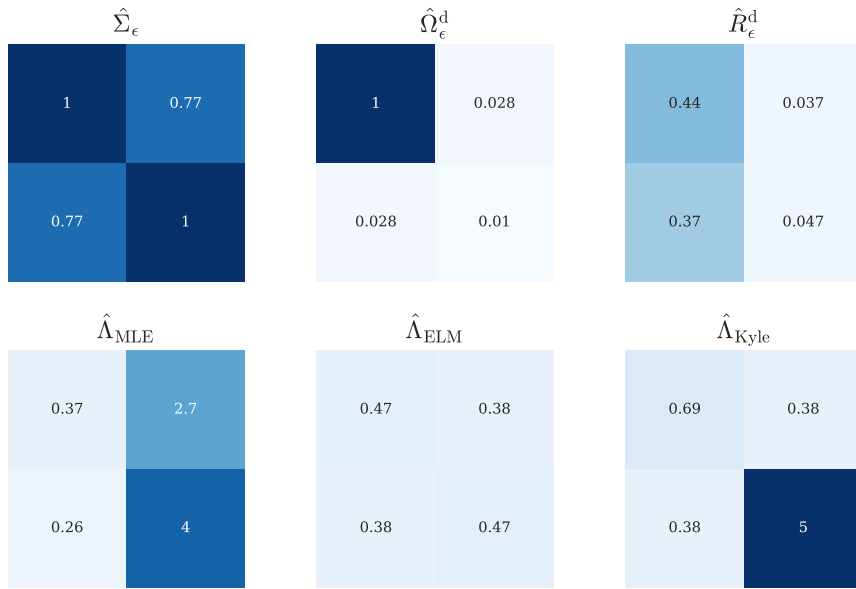


Figure 3: **Extreme illiquidity.** Covariance matrices and estimators with an illiquid asset. The value of $\rho = 0.77$ is the empirical correlation of the representative pair 2USNOTES-5USNOTES, whereas the value of $\epsilon = 0.01$ is obtained by rescaling the empirical responses and volume covariances of the same pair.

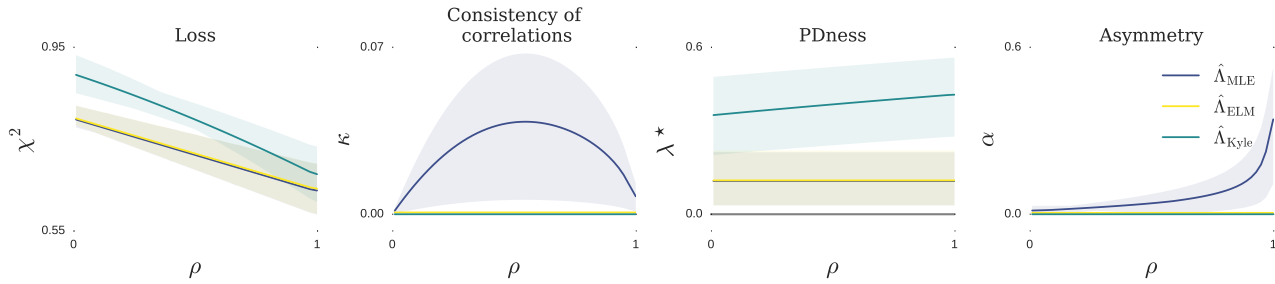


Figure 4: **Effect of price variation correlations** Averaged values (solid lines) and intervals between the maximum and the minimum (shaded area) of the observables for ρ ranging from 0 to 1.

- $\hat{\Lambda}_{Kyle}$: The main difference between the predictions of the Kyle estimator and $\hat{\Lambda}_{MLE}$ is that in this case trading the second asset will strongly modify its price but it will have very little impact on the first assets' price. As already emphasized, this is how efficient markets should work: liquid instruments should not be affected by the order flow on correlated illiquid instruments, otherwise price manipulation strategies would be possible.

Extreme correlations We now look at a linear combinations of two of the assets with price variation covariances:

$$\hat{\Sigma} \mapsto \hat{\Sigma}_\rho = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} .$$

In Fig. 4 we show the value of each of the observables averaged over the 6 sets of covariances for ϵ ranging from 0 to 1. The loss of all estimators decreases as $\rho \rightarrow 1$ and as in the previous case, the distributions have a

strong overlap, suggesting again that no one estimator is significantly better at minimizing the loss.

By construction $\kappa_{\text{Kyle}}=0$, and in this case $\kappa_{\text{ELM}} = 0$ because $\hat{\Omega}_\rho^d$ commutes with $\hat{\Sigma}_\rho$ (it would not be the case if the volume variances were not the same). Also $\kappa_{\text{MLE}} = 0$ for $\rho = 0$ because in that case $\hat{\Sigma}_\rho = \mathbb{I}$, but for $\rho > 0$, $\kappa_{\text{MLE}} > 0$. The distributions of λ_{ELM}^* and λ_{MLE}^* are essentially the same. λ_{Kyle}^* is significantly larger than the other two for all values of ρ . Asymmetries in $\hat{\Lambda}_{\text{MLE}}$ are amplified as $\rho \rightarrow 1$ because in that limit $\hat{\Omega}_\rho^d$ becomes singular.

Again, following example 7 in Fig. 5 we show the covariance matrices and the three corresponding impact estimators for a case where the price variations of the two assets are strongly correlated. Let us analyze the estimators obtained:

- $\hat{\Lambda}_{\text{MLE}}$: Small arbitrary fluctuations in \hat{R}_ρ^d are amplified due to the strong volume correlations that result from the transformation. In the example shown in the figure the impact on the second asset is much larger than on the first one.
- $\hat{\Lambda}_{\text{ELM}}$: Self impact and cross-impact are comparable. Furthermore, self-impact is very similar for both assets.
- $\hat{\Lambda}_{\text{Kyle}}$: Self-impact is very similar for both assets and stronger than cross-impact.

6 Conclusion

In the present work we studied a Multivariate Kyle model, that proved to be a very interesting setting to understand the fundamental mechanisms underlying *cross-impact*. Perhaps more importantly, the Multivariate Kyle model suggests a practical recipe to extract a consistent cross-impact matrix structure from empirical data – a point that does not seem to have been emphasized before but that becomes crucial when dealing with present day large dimensional data. We recovered the Caballé-Krishnan solution at equilibrium and proved the unicity of the symmetric solution. We provided an interpretation of the results with regard to the eigen-modes of returns and volumes covariances. We discussed the implications of the model for pricing, with a particular focus on the SPD property of the impact matrix. We presented the implications for cross-impact regression of empirical data, and showed that cross-awareness (or in lesser terms “I know that you know”) can be used as powerful regulariser with small loss in predictive power. We confronted our results with previous empirical cross-impact analyses, and identified limiting regimes in which the results in [BMEB17, MBEB17] are reproduced.

From a complementary point of view, our results can be seen as proxies for the behavior of an idealized market in which prices fully reflect the information encoded in the order flow: for a manipulation-free market, prices of liquid securities should be insensitive to trading of strongly illiquid instruments (see example 6), whereas prices of strongly correlated instruments should be insensitive to how they are individually traded (example 7). Measuring how much real markets violate these principles would be an interesting empirical application of our results, that could be used to assist regulators in order to assess the vulnerability of a market to correlated trading activity.

In spite of all the “good” properties of the Kyle cross-impact estimator, one should keep in mind that many important aspects are left out of the Kyle framework, that may play a crucial role in practice. First of all, the empirical order flow is strongly autocorrelated in time, which must induce a non-trivial lag dependence of the impact function – as found in [BMEB17]. Extending the present theory to predict a lag-dependent impact matrix $\Lambda(\tau)$ would be extremely useful. Second, impact is non-linear in traded quantities: it is now well accepted that the impact of a metaorder has a square-root dependence on volume. How this single asset square-root law generalizes in a multiasset context is, as far as we are aware, a completely open problem. Finally, it would be interesting to extend to the multivariate case some recent extensions of the Kyle model that accounts for the inventory risk of the market maker [ÇD⁺16]. We hope to visit some of these questions in future work.

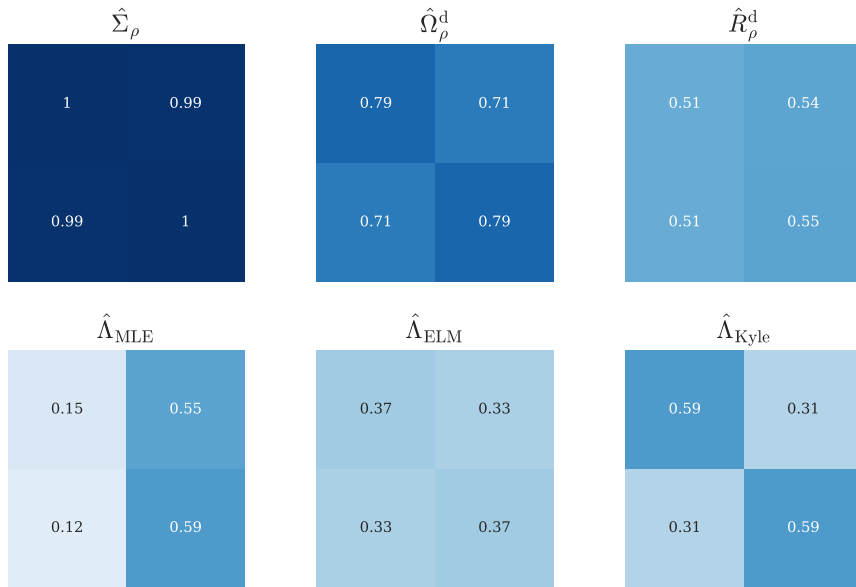


Figure 5: **Strong price variation correlations.** Covariance matrices and estimators with strongly correlated price variations ($\rho = 0.99$)

Acknowledgments

We thank Z. Eisler, A. Fosset and E. Sérié for very fruitful discussions. We also thank J. Caballé for kindly providing the unpublished manuscript [CK⁺90].

References

- [AKS16] Aurélien Alfonsi, Florian Klöck, and Alexander Schied. Multivariate transient price impact and matrix-valued positive definite functions. *Mathematics of operations research*, 41(3):914–934, 2016.
- [BBDG18] Jean-Philippe Bouchaud, Julius Bonart, Jonathan Donier, and Martin Gould. *Trades, quotes and prices: financial markets under the microscope*. Cambridge University Press, 2018.
- [BMEB17] Michael Benzaquen, Iacopo Mastromatteo, Zoltan Eisler, and Jean-Philippe Bouchaud. Dissecting cross-impact on stock markets: An empirical analysis. *Journal of Statistical Mechanics: Theory and Experiment*, 2017(2):023406, 2017.
- [ÇD⁺16] Umut Çetin, Albina Danilova, et al. Markovian nash equilibrium in financial markets with asymmetric information and related forward–backward systems. *The Annals of Applied Probability*, 26(4):1996–2029, 2016.
- [CK⁺90] Jordi Caballé, Murugappa Krishnan, et al. Insider trading and asset pricing in an imperfectly competitive multi-security market. Technical report, Unitat de Fonaments de l’Anàlisi Econòmica (UAB) and Institut d’Anàlisi Econòmica (CSIC), 1990.
- [CK94] Jordi Caballe and Murugappa Krishnan. Imperfect competition in a multi-security market with risk neutrality. *Econometrica: Journal of the Econometric Society*, pages 695–704, 1994.
- [HJ85] Roger A Horn and Charles R. Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, 1985.
- [HS01] Joel Hasbrouck and Duane J Seppi. Common factors in prices, order flows, and liquidity. *Journal of financial Economics*, 59(3):383–411, 2001.
- [Kyl85] Albert S Kyle. Continuous auctions and insider trading. *Econometrica: Journal of the Econometric Society*, pages 1315–1335, 1985.
- [Las04] Guillaume Lasserre. Asymmetric information and imperfect competition in a continuous time multivariate security model. *Finance and Stochastics*, 8(2):285–309, 2004.
- [LRS13] Georg Lindgren, Holger Rootzén, and Maria Sandsten. *Stationary stochastic processes for scientists and engineers*. CRC press, 2013.
- [MBEB17] Iacopo Mastromatteo, Michael Benzaquen, Zoltan Eisler, and Jean-Philippe Bouchaud. Trading lightly: Cross-impact and optimal portfolio execution. *Risk Magazine*, July 2017:1–6, 2017.
- [PV13] Paolo Pasquariello and Clara Vega. Strategic cross-trading in the us stock market. *Review of Finance*, 19(1):229–282, 2013.
- [SL18] Michael Schneider and Fabrizio Lillo. Cross-impact and no-dynamic-arbitrage. *Quantitative Finance*, pages 1–18, 2018.
- [Vit12] Paolo Vitale. Risk-averse insider trading in multi-asset sequential auction markets. *Economics Letters*, 117(3):673–675, 2012.
- [WSG16] Shanshan Wang, Rudi Schäfer, and Thomas Guhr. Cross-response in correlated financial markets: individual stocks. *The European Physical Journal B*, 89(4):105, 2016.

A Solution of the Multivariate Kyle model

The proof of Theorem 1 is split in 4 different results. First, the fact that the equilibrium is linear is a consequence of Propositions 1 and 2. Before finding the explicit form of the equilibrium we need to show that $\mathbf{\Lambda}$ is symmetric. This is a consequence of the second order condition on the minimization of the IT's strategy (Proposition 1) and is proven in Proposition 3. Finally, the explicit form and proof of uniqueness of $\mathbf{\Lambda}$ is given in Proposition 11

A.1 Optimality of the Informed Trader

Proof of Proposition 1. Here, we show the consequences of the first and second order conditions arising from the requirement that the IT is maximizing locally the average utility given \mathbf{v} , which can be written as:

$$\begin{aligned}\mathbb{E}[\mathcal{U}_{IT}|\mathbf{v}] &= -\mathbb{E}[\mathbf{x}^\top(\mathbf{p} - \mathbf{v})|\mathbf{v}] \\ &= -\mathbb{E}[\mathbf{x}^\top(\boldsymbol{\mu} + \mathbf{\Lambda}\mathbf{y} - \mathbf{v})|\mathbf{v}] \\ &= -\mathbf{x}^\top\boldsymbol{\mu} - \mathbf{x}^\top\mathbf{\Lambda}\mathbf{x} + \mathbf{x}^\top\mathbf{v}.\end{aligned}$$

In order to find the optimal strategy, we maximize the utility with respect to \mathbf{x} . The first and second order derivatives are:

$$\frac{\partial\mathbb{E}[\mathcal{U}_{IT}|\mathbf{v}]}{\partial\mathbf{x}} = -\boldsymbol{\mu} - 2\mathbf{\Lambda}_S\mathbf{x} + \mathbf{v} = 0 \quad (60)$$

$$\frac{\partial^2\mathbb{E}[\mathcal{U}_{IT}|\mathbf{v}]}{\partial\mathbf{x}\partial\mathbf{x}^\top} = -2\mathbf{\Lambda}_S. \quad (61)$$

where \mathbf{X}_S (\mathbf{X}_A) denotes the symmetric (antisymmetric) part of \mathbf{X} .

Since \mathcal{U}_{IT} is quadratic on \mathbf{x} , the profit maximization condition implies that the second derivative (61) has to be strictly negative definite, which in turn implies that both $\mathbf{\Lambda}$ and $\mathbf{\Lambda}_S$ have to be PD. Solving (60) for \mathbf{x} leads to

$$\mathbf{x} = \boldsymbol{\alpha} + \mathbf{B}\mathbf{v} \quad (62)$$

where

$$\begin{aligned}\boldsymbol{\alpha} &= -\mathbf{B}\boldsymbol{\mu} \\ \mathbf{B} &= \frac{1}{2}\mathbf{\Lambda}_S^{-1}.\end{aligned} \quad (63)$$

Since $\mathbf{\Lambda}_S$ has to be PD, it is invertible and \mathbf{B} and $\boldsymbol{\alpha}$ are well defined. \square

As we will show in App. A.3, PDness of $\mathbf{\Lambda}$ together with the efficient-price condition (discussed in App. A.2) implies that $\mathbf{\Lambda}$ is symmetric.

A.2 Optimality of the Market Maker

In this Appendix we will state the conditions stemming from the requirement that the MM fixes a pricing rule $\mathbf{p} = \mathbf{\Lambda}\mathbf{y} + \boldsymbol{\mu}$ such that $\mathbf{p} = \mathbb{E}[\mathbf{v}|\mathbf{y}]$ is an efficient price.

Proof of Proposition 2. Exploiting the Gaussian nature of \mathbf{v} and \mathbf{u} and their independence, the efficient price as given in (2) can be expanded as [LRS13, p. 269]:

$$\begin{aligned}\mathbf{p} = \mathbb{E}[\mathbf{v}|\mathbf{y}] &= \mathbb{E}[\mathbf{v}] + \mathbf{C}[\mathbf{v}, \mathbf{y}]\mathbf{C}[\mathbf{y}, \mathbf{y}]^{-1}(\mathbf{y} - \mathbb{E}[\mathbf{y}]) \\ &= \mathbf{p}_0 + \mathbf{R}_v^d(\boldsymbol{\Omega}^d)^{-1}(\mathbf{y} - \mathbf{y}_0) \\ &= \boldsymbol{\mu} + \mathbf{\Lambda}\mathbf{y},\end{aligned} \quad (64)$$

where

$$\boldsymbol{\mu} = \boldsymbol{p}_0 - \boldsymbol{\Lambda} \boldsymbol{y}_0, \quad (65)$$

$$\boldsymbol{\Lambda} = \boldsymbol{R}_v^d (\boldsymbol{\Omega}^d)^{-1}. \quad (66)$$

□

This analysis provides an explicit form for the MM's strategy and it might look as if the MM is doing a linear regression. However, without knowing the strategy of the IT one cannot estimate \boldsymbol{R}_v^d nor $(\boldsymbol{\Omega}^d)^{-1}$ meaning that the MM has to include information about \boldsymbol{x} in his regression making it a quadratic problem rather than a linear one. This point is discussed in more detail in Sec. 5.2.1.

A.3 Symmetry of the linear equilibrium

In order to find the linear equilibrium described in Theorem 1, we first have to prove the proposition 3. To do so we will make use of an useful lemma [HJ85, Theorem 7.2.6] that will be used in the following derivations.

Lemma 1 (Positive definite square-root). *Consider a matrix \boldsymbol{Y} that is symmetric and positive semi-definite. Then, it exists a unique symmetric, positive semi-definite matrix $\boldsymbol{X} = \sqrt{\boldsymbol{Y}}$ such that $\boldsymbol{X}\boldsymbol{X} = \boldsymbol{Y}$.*

Proof of Proposition 3. Let us start by plugging $\mathbb{E}[\boldsymbol{y}] = \boldsymbol{y}_0 = \frac{1}{2} \boldsymbol{\Lambda}_S^{-1} (\boldsymbol{p}_0 - \boldsymbol{\mu})$ into (15). After a bit of algebra we obtain

$$(\boldsymbol{p}_0 - \boldsymbol{\mu}) \left(\mathbb{I} - \frac{1}{2} \boldsymbol{\Lambda} \boldsymbol{\Lambda}_S^{-1} \right) = 0 \quad (67)$$

thus $\boldsymbol{\mu} = \boldsymbol{p}_0$ and $\boldsymbol{x} = \frac{1}{2} \boldsymbol{\Lambda}_S^{-1} (\boldsymbol{v} - \boldsymbol{p}_0)$ which allows us to compute the following quantities:

$$\boldsymbol{y}_0 = 0 \quad (68)$$

$$\boldsymbol{\Omega}^d = \frac{1}{4} \boldsymbol{\Lambda}_S^{-1} \boldsymbol{\Sigma}_0 \boldsymbol{\Lambda}_S^{-1} + \boldsymbol{\Omega} \quad (69)$$

$$\boldsymbol{R}_v^d = \frac{1}{2} \boldsymbol{\Sigma}_0 \boldsymbol{\Lambda}_S^{-1}. \quad (70)$$

By plugging the results above into (15) one obtains a quadratic equation:

$$(\boldsymbol{\Lambda} \boldsymbol{\Lambda}_S^{-1} - 2\mathbb{I}) \frac{1}{4} \boldsymbol{\Sigma}_0 + \boldsymbol{\Lambda} \boldsymbol{\Omega} \boldsymbol{\Lambda}_S = 0. \quad (71)$$

Expanding $\boldsymbol{\Lambda}$ as $(\boldsymbol{\Lambda}_A + \boldsymbol{\Lambda}_S)$ and doing some algebra from (71) we obtain

$$\boldsymbol{\Lambda}_A = (\mathbb{I} - \boldsymbol{\Gamma})(\mathbb{I} + \boldsymbol{\Gamma})^{-1} \boldsymbol{\Lambda}_S \quad (72)$$

where $\boldsymbol{\Gamma} = 4\boldsymbol{\Lambda}_S \boldsymbol{\Omega} \boldsymbol{\Lambda}_S \boldsymbol{\Sigma}_0^{-1}$. Now, since $\boldsymbol{\Lambda}_A = \frac{1}{2}(\boldsymbol{\Lambda} - \boldsymbol{\Lambda}^\top)$, we have $\boldsymbol{\Lambda}_A = -\boldsymbol{\Lambda}_A^\top$ and therefore

$$\begin{aligned} (\mathbb{I} - \boldsymbol{\Gamma})(\mathbb{I} + \boldsymbol{\Gamma})^{-1} \boldsymbol{\Lambda}_S &= -\boldsymbol{\Lambda}_S (\mathbb{I} + \boldsymbol{\Gamma}^\top)^{-1} (\mathbb{I} - \boldsymbol{\Gamma}^\top) \\ (\mathbb{I} + \boldsymbol{\Gamma}^\top) \boldsymbol{\Lambda}_S^{-1} (\mathbb{I} - \boldsymbol{\Gamma}) &= -(\mathbb{I} - \boldsymbol{\Gamma}^\top) \boldsymbol{\Lambda}_S^{-1} (\mathbb{I} + \boldsymbol{\Gamma}) \\ \sqrt{\boldsymbol{\Lambda}_S} \boldsymbol{\Gamma}^\top \boldsymbol{\Lambda}_S^{-1} \boldsymbol{\Gamma} \sqrt{\boldsymbol{\Lambda}_S} &= \mathbb{I}. \end{aligned}$$

Substituting back $\boldsymbol{\Gamma}$ we get

$$\boldsymbol{Y}^2 = \boldsymbol{Z}^2$$

where $\boldsymbol{Y} = \sqrt{\boldsymbol{\Lambda}_S} \boldsymbol{\Omega} \sqrt{\boldsymbol{\Lambda}_S}$ and $\boldsymbol{Z} = \sqrt{\boldsymbol{\Lambda}_S^{-1} \frac{1}{4} \boldsymbol{\Sigma}_0 \boldsymbol{\Lambda}_S^{-1}}$. Since $\boldsymbol{\Lambda}_S$ is PD, \boldsymbol{Y}^2 is also SPD and therefore $\boldsymbol{Y} = \boldsymbol{Z}$ is the unique PD square root of \boldsymbol{Y}^2 which leads to

$$\frac{1}{4} \boldsymbol{\Sigma}_0 = \boldsymbol{\Lambda}_S \boldsymbol{\Omega} \boldsymbol{\Lambda}_S \quad (73)$$

and substituting these results into (72) reveals that $\mathbf{\Lambda}_A = 0$. \square

With equation (73) we prove that the profit maximization condition on the IT's cost (i.e. $\mathbf{\Lambda}$ has to be PD) results in $\mathbf{\Lambda}_A = 0$ and $\mathbf{\Lambda} = \mathbf{\Lambda}_S$. If $\mathbf{\Lambda}_S$ was not PD, \mathbf{Y} and \mathbf{Z} could be different roots and we would not be able to establish the relationship $\mathbf{Y} = \mathbf{Z}$. In that case we would find many solutions, most of them being saddle points for $\mathbb{E}[\mathcal{U}_{IT}]$.

A.4 Explicit form of the linear equilibrium

The last step needed in order to prove the existence and uniqueness the linear equilibrium in Theorem 1 is to prove that Eq. (73) admits a unique symmetric solution, as made explicit below.

Proposition 11. *The unique symmetric PD matrix $\mathbf{\Lambda}$ satisfying Eq. (73) is*

$$\mathbf{\Lambda} = \frac{1}{2} \mathcal{R}^{-1} \sqrt{\mathcal{R} \mathbf{\Sigma}_0 \mathcal{L} \mathcal{L}^{-1}}.$$

Proof. Explicit form of $\mathbf{\Lambda}$: The solution to Eq. (73) has to be a $\mathbf{\Lambda}$ of the form

$$\mathbf{\Lambda} = \frac{1}{2} \mathcal{G} \mathbf{O} \mathcal{L}^{-1}, \quad (74)$$

where \mathcal{G} is any factorization of $\mathbf{\Sigma}_0$ of the form $\mathcal{G} \mathcal{D}$ with $\mathcal{G}^\top = \mathcal{D}$. One is then left with the equation $\mathbf{O} \mathbf{O}^\top = \mathbb{I}$. Such equation is solved by *any* matrix \mathbf{O} belonging to the orthogonal group $O(n)$, which is specified by $n(n-1)/2$ parameters. It is only by self-consistently imposing symmetry of $\mathbf{\Lambda}$ that one finds the solution for \mathbf{O}

$$\begin{aligned} \mathbf{O} &= \mathcal{G}^{-1} \mathcal{R}^{-1} \sqrt{\mathcal{R} \mathbf{\Sigma}_0 \mathcal{L}} \\ &= (\mathcal{R} \mathcal{G})^{-1} \sqrt{(\mathcal{R} \mathcal{G})(\mathcal{R} \mathcal{G})^\top}. \end{aligned} \quad (75)$$

Introducing this value of \mathbf{O} in Eq. (74) we obtain the desired result.

Uniqueness: The decomposition $\mathbf{\Omega} = \mathcal{L} \mathcal{R}$ is not unique. Indeed, multiplying \mathcal{L} by an arbitrary rotation \mathbf{O}_Ω yields another possible decomposition. However the final value of $\mathbf{\Lambda}$ does not depend on \mathbf{O}_Ω . To prove this consider the value of $\mathbf{\Lambda}$ using the decomposition $\mathbf{\Omega} = \mathcal{L} \mathbf{O}_\Omega \mathbf{O}_\Omega^\top \mathcal{R}$:

$$\begin{aligned} \mathbf{\Lambda} &= \frac{1}{2} (\mathbf{O}_\Omega^\top \mathcal{R})^{-1} \sqrt{\mathbf{O}_\Omega^\top \mathcal{R} \mathbf{\Sigma}_0 \mathcal{L} \mathbf{O}_\Omega (\mathcal{L} \mathbf{O}_\Omega)^{-1}} \\ &= \frac{1}{2} \mathcal{R}^{-1} \mathbf{O}_\Omega \mathbf{O}_\Omega^\top \sqrt{\mathcal{R} \mathbf{\Sigma}_0 \mathcal{L} \mathbf{O}_\Omega \mathbf{O}_\Omega^\top \mathcal{L}^{-1}} \\ &= \frac{1}{2} \mathcal{R}^{-1} \sqrt{\mathcal{R} \mathbf{\Sigma}_0 \mathcal{L} \mathcal{L}^{-1}} \end{aligned}$$

where we have used the fact that if $\sqrt{\mathbf{Y}} = \mathbf{X}$ then $\sqrt{\mathbf{O}^\top \mathbf{Y} \mathbf{O}} = \mathbf{O}^\top \mathbf{X} \mathbf{O}$. Since the argument of the square root is SPD by construction the unique value of $\mathbf{\Lambda}$ is obtained by choosing the unique SPD root. \square

B Estimation

B.1 The loss χ^2

First, we want to relate the loss (48) with the empirical mean and covariances defined in Sec. 5. Starting from the definition (48), one can write

$$\begin{aligned} \chi^2 &= \frac{1}{2} \langle (\hat{\mathbf{\Lambda}} \mathbf{y} - \Delta \mathbf{p})^\top \mathbf{M} (\hat{\mathbf{\Lambda}} \mathbf{y} - \Delta \mathbf{p}) \rangle, \\ &= \frac{1}{2} \langle \mathbf{y}^\top \hat{\mathbf{\Lambda}}^\top \mathbf{M} \hat{\mathbf{\Lambda}} \mathbf{y} \rangle - \langle \mathbf{y}^\top \hat{\mathbf{\Lambda}}^\top \mathbf{M} \Delta \mathbf{p} \rangle + \frac{1}{2} \langle \Delta \mathbf{p}^\top \mathbf{M} \Delta \mathbf{p} \rangle \end{aligned}$$

Then, by plugging the definitions of the empirical covariances one is left with

$$\chi^2 = \frac{1}{2} \text{tr} \left[\hat{\mathbf{\Lambda}}^\top \mathbf{M} \hat{\mathbf{\Lambda}} \hat{\mathbf{\Omega}}^d - 2 \hat{\mathbf{\Lambda}}^\top \mathbf{M} \hat{\mathbf{R}}^d + \mathbf{M} \hat{\mathbf{\Sigma}} \right] \quad (76)$$

B.2 Maximum Likelihood Estimator

Proof of Proposition 8. In order to derive the Maximum Likelihood estimator, we differentiate (76) with respect to $\hat{\mathbf{\Lambda}}$:

$$0 = \frac{\partial \chi^2}{\partial \hat{\mathbf{\Lambda}}} = \mathbf{M} (\hat{\mathbf{\Lambda}} \hat{\mathbf{\Omega}}^d - \hat{\mathbf{R}}^d).$$

It is clear that the optimal value of $\hat{\mathbf{\Lambda}}_{\text{MLE}}$ does not depend on \mathbf{M} . To find the explicit expression we equate the derivative to 0 and solve:

$$\hat{\mathbf{\Lambda}}_{\text{MLE}} = \hat{\mathbf{R}}^d (\hat{\mathbf{\Omega}}^d)^{-1}.$$

To calculate the minimum we can plug this solution into Eq. (76), obtaining

$$\chi_{\text{MLE}}^2 = \frac{1}{2} \text{tr} \left[\mathbf{M} \left(\hat{\mathbf{\Sigma}} - \hat{\mathbf{R}}^d (\hat{\mathbf{\Omega}}^d)^{-1} (\hat{\mathbf{R}}^d)^\top \right) \right]. \quad (77)$$

□

B.3 EigenLiquidity Estimator

The derivation of the EigenLiquidity Estimator runs along the same lines as the previous case.

Proof of Proposition 9. We want to find an estimator $\hat{\mathbf{\Lambda}}_{\text{ELM}}$ that minimizes χ^2 under the constraint $[\hat{\mathbf{\Lambda}}_{\text{ELM}}, \hat{\mathbf{\Sigma}}] = 0$ which is equivalent to saying that

$$\hat{\mathbf{\Lambda}}_{\text{ELM}} = \sum_{a=1}^n \hat{\mathbf{s}}_a g_a \hat{\mathbf{s}}_a^\top$$

where $\{\hat{\mathbf{s}}_a\}_{a=1}^n$ are the eigenvectors of $\hat{\mathbf{\Sigma}}$. Because of this constraint, the minimization of χ^2 has to be done with respect to g_a rather than $\hat{\mathbf{\Lambda}}$:

$$0 = \frac{\partial \chi^2}{\partial g_a} = \text{tr} \left(\mathbf{M} (\hat{\mathbf{\Lambda}} \hat{\mathbf{\Omega}}^d - \hat{\mathbf{R}}^d) \hat{\mathbf{s}}_a \hat{\mathbf{s}}_a^\top \right),$$

which for arbitrary \mathbf{M} implies

$$\hat{\mathbf{s}}_a^\top (\hat{\mathbf{\Lambda}} \hat{\mathbf{\Omega}}^d - \hat{\mathbf{R}}^d) \hat{\mathbf{s}}_a = 0$$

and therefore

$$g_a = \frac{\hat{\mathbf{s}}_a^\top \hat{\mathbf{R}}^d \hat{\mathbf{s}}_a}{\hat{\mathbf{s}}_a^\top \hat{\mathbf{\Omega}}^d \hat{\mathbf{s}}_a}.$$

In order to compute the minimum loss we proceed as in the previous case plugging the value of $\hat{\mathbf{\Lambda}}_{\text{ELM}}$ into equation (76) and obtain

$$\begin{aligned} \chi_{\text{ELM}}^2 &= \frac{1}{2} \text{tr} \left[\mathbf{M} \left(\hat{\mathbf{\Sigma}} + (\hat{\mathbf{\Lambda}}_{\text{ELM}} \hat{\mathbf{\Omega}}^d - 2 \hat{\mathbf{R}}^d) \hat{\mathbf{\Lambda}}_{\text{ELM}} \right) \right] \\ &= \frac{1}{2} \text{tr} \left[\mathbf{M} \left(\hat{\mathbf{\Sigma}} - \hat{\mathbf{R}}^d \hat{\mathbf{\Lambda}}_{\text{ELM}} \right) \right] \\ &= \frac{1}{2} \text{tr} \left[\mathbf{M} \left(\hat{\mathbf{\Sigma}} - \hat{\mathbf{R}}^d \sum_{a=1}^n \hat{\mathbf{s}}_a \frac{\hat{\mathbf{s}}_a^\top \hat{\mathbf{R}}^d \hat{\mathbf{s}}_a}{\hat{\mathbf{s}}_a^\top \hat{\mathbf{\Omega}}^d \hat{\mathbf{s}}_a} \hat{\mathbf{s}}_a^\top \right) \right]. \end{aligned}$$

□

B.4 Kyle Estimator

Proof of Proposition 10 To prove the first statement of the proposition it suffices to compute the covariance of the efficient prices:

$$\hat{\Sigma}_{\text{Kyle}} = \langle \hat{\Delta p} \hat{\Delta p}^\top \rangle = \hat{\Lambda}_{\text{Kyle}} \langle \mathbf{y} \mathbf{y}^\top \rangle \hat{\Lambda}_{\text{Kyle}}^\top = \hat{\Lambda}_{\text{Kyle}} \hat{\Omega}^d \hat{\Lambda}_{\text{Kyle}}^\top = k^2 \hat{\Sigma},$$

and to use the unicity result for the above equation (proved in App. A.4) in order to recover Eq. (58).

In order to find the loss at the minimum one has to take into account that the only free parameter is k , and since $\hat{\Lambda}_{\text{Kyle}} \hat{\Omega}^d \hat{\Lambda}_{\text{Kyle}}^\top = k^2 \hat{\Sigma}$, the loss is

$$\begin{aligned} \chi_{\text{Kyle}}^2 &= \frac{1}{2} \text{tr} \left[M \left(\hat{\Sigma} + (\hat{\Lambda}_{\text{ELM}} \hat{\Omega}^d - 2\hat{R}^d) \hat{\Lambda}_{\text{Kyle}} \right) \right] \\ &= \frac{1}{2} \text{tr} \left[M \left((1+k^2) \hat{\Sigma} - 2k \hat{R}^d (\hat{\mathcal{R}}^d)^{-1} \sqrt{\hat{\mathcal{R}}^d \hat{\Sigma} \hat{\mathcal{L}}^d (\hat{\mathcal{L}}^d)^{-1}} \right) \right]. \end{aligned}$$

Optimizing with respect to k we obtain

$$\frac{\partial \chi_{\text{Kyle}}^2}{\partial k} = k \text{tr} \left[M \hat{\Sigma} \right] - \text{tr} \left[M \hat{R}^d (\hat{\mathcal{R}}^d)^{-1} \sqrt{\hat{\mathcal{R}}^d \hat{\Sigma} \hat{\mathcal{L}}^d (\hat{\mathcal{L}}^d)^{-1}} \right] = 0.$$

□

B.5 Fabricating synthetic covariance matrices inspired from real data

The data set used to fabricate the synthetic covariance matrices in section 5.3 consists of the averaged price variations and volumes in 5 minute bins of the 2-year, 5-year, 10-year, and 30-year tenors of the U.S. Treasury Futures traded in the Chicago Board of Trade during the year 2016.

For each pair of bonds with price variations $\Delta \mathbf{p}^{(t)}$ and volumes $\mathbf{y}^{(t)}$ we define the normalized covariances as

$$\mathbf{C} = \begin{pmatrix} \hat{\Sigma} & (\hat{R}^d)^\top \\ \hat{R}^d & \hat{\Omega}^d \end{pmatrix} = \mathbf{D}^{-1} \begin{pmatrix} \langle \Delta \mathbf{p} \Delta \mathbf{p}^\top \rangle & \langle \mathbf{y} \Delta \mathbf{p}^\top \rangle \\ \langle \Delta \mathbf{p} \mathbf{y}^\top \rangle & \langle \mathbf{y} \mathbf{y}^\top \rangle \end{pmatrix} \mathbf{D}^{-1}$$

where we have previously shifted the data by their empirical means and where

$$\mathbf{D} = \text{diag} \left(\sqrt{\langle \Delta p_1^2 \rangle}, \sqrt{\langle \Delta p_2^2 \rangle}, \sqrt{\langle y_1^2 \rangle}, \sqrt{\langle y_2^2 \rangle} \right).$$

Fixing liquidity: Covariance matrices with the desired liquidity are constructed by rescaling the rows and columns of \mathbf{C}

$$\mathbf{C}_\epsilon = \begin{pmatrix} \hat{\Sigma}_\epsilon & (\hat{R}_\epsilon^d)^\top \\ \hat{R}_\epsilon^d & \hat{\Omega}_\epsilon^d \end{pmatrix} = \begin{pmatrix} 1 & & 0 \\ & 1 & \\ 0 & & 1 \\ & & & \sqrt{\epsilon} \end{pmatrix} \mathbf{C} \begin{pmatrix} 1 & & 0 \\ & 1 & \\ 0 & & 1 \\ & & & \sqrt{\epsilon} \end{pmatrix}.$$

Using this recipe, the variance of the volume of the second asset will be ϵ and consequently the second column of the response is multiplied by $\sqrt{\epsilon}$. Note that the price variations are unchanged.

Fixing correlations: In order to modify the price correlations we have to construct linear combinations of assets. If $\hat{\Sigma}_{12} = r$, then

$$\mathbf{C}_\rho = \begin{pmatrix} \hat{\Sigma}_\rho & (\hat{R}_\rho^d)^\top \\ \hat{R}_\rho^d & \hat{\Omega}_\rho^d \end{pmatrix} = \begin{pmatrix} \mathbf{A} & 0 \\ 0 & \mathbf{A} \end{pmatrix} \mathbf{C} \begin{pmatrix} \mathbf{A} & 0 \\ 0 & \mathbf{A} \end{pmatrix}$$

where

$$\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} \sqrt{\frac{1+\rho}{1+r}} & 0 \\ 0 & \sqrt{\frac{1-\rho}{1-r}} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

This modification ensures that $\hat{\Sigma}_\rho = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$. The response and volumes are also modified accordingly. In particular the diagonal entries of $\hat{\Omega}_\rho^d$ will still be equal (but not necessarily equal to 1).

Note that for \mathbf{C} SPD, both \mathbf{C}_ϵ and \mathbf{C}_ρ are also SPD and therefore are coherent covariance matrices.